Spatially Index-based Assessment of Flash Flood Potential under Sparse Data Availability

Von der
Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig

zu Erlangung des Grades eines
Doktoringenieurs (Dr.-Ing.)
genehmigte

Dissertation

von
Worapong Lohpaisankrit
geboren am 05.06.1980
aus Khon Kaen, Thailand

Eingereicht am:	03. Mai 2018
Disputation am:	28. Juni 2018

Berichterstatter:	Prof. Dr.-Ing. Günter Meon Prof. Dr. Tawatchai Tingsanchali

2018

Suggested citation:

Lohpaisankrit, W. (2018): Spatially Index-based Assessment of Flash Flood Potential under Sparse Data Availability. PhD dissertation, Leichtweiss Institute for Hydraulic Engineering and Water Resources, Technische Universitaet Braunschweig, Braunschweig, Germany.

This work is dedicated to my parents, Saowapa and Pongsak Lohpaisankrit,
my dear sister, Wannaporn Lohpaisankrit Phipps,
my lovely wife Piyarut, and son Bordinpath who have supported me
in many ways throughout this journey.

Acknowledgments

First and foremost, I would like to express my sincere appreciation to my first supervisor, Prof. Dr.-Ing. Günter Meon, for initiating my ideas for this PhD research and patiently guiding throughout last four years of my studies. Without his guidance, support and inspiration, many research difficulties would not have been defeated. He has not only provided friendly working environments but also handed me substantial opportunities that have strengthened my research skills and independence.

Secondly, I am especially indebted to my second supervisor, Prof. Dr. Tawatchai Tingsanchali, who introduced me to Prof. Dr.-Ing. Meon. His connection has started my PhD journey. Therefore, I would like to take this opportunity to deliver my heartfelt thanks to Prof. Dr. Tingsanchali. His excellent knowledge, experience and thoughtful comments helped me to deploy great ideas and to scientifically develop this dissertation.

Thirdly, I am particularly grateful to Dr.-Ing. Phillip Kreye for his valuable suggestions and discussions. Also, thanks to his lovely family for providing some support and useful guidelines to my family.

Fourthly, I would like to gratefully thank to Dr.-Ing. Michael Leismann and Kai Otte-Witte, Ingenieure für Wasser, Umwelt und Datenverarbeitung (IWUD) GmbH, for cost-free provision of the MEADFLOW software and for technical support.

Additionally, I would like to extend my earnest appreciation to the staff members and my colleagues at the Leichtweiss Institute for Hydraulic Engineering and Water Resources (LWI) – namely, Prof. Dr. Matthias Schöniger, Dr.-Ing. Gerhard Riedel, Dr.-Ing. Huyen Le, Dr.-Ing. Malte Lorenz, Dr.-Ing. Tien Thanh Nguyen, Dr.-Ing. Saber M. Elsayed, Susanne Festerling, Ehna Dümpert, Malte Eley, Marlene Gelleszun, Tim Lichtenberg, Tim Müller, Hernita Nasir, Saskia Schimmelpfennig, Manickam Somasundaram, Karoline Stein, Le Minh Vu, Vanessa Wörner, and Stephanie Zeunert, for good advice and cooperation as well as good friendships.

I would like to acknowledge many MATLAB and R communities for sharing codes and discussing many useful techniques in programming on the accessible Internet platforms.

Needless to say, this undertaking has required sacrifice on the part of my parents, my dear sister Wannaporn, my son Bordinpath, and especially my wife Piyarut. A great deal of credit must share with them. Without their love and support, none of this work would have been possible.

Last but not least, I would like to give big thanks to the German Academic Exchange Service (DAAD[1]) and Technische Universität Braunschweig for financial support. Many thanks go to the Royal Irrigation Department (RID), Thai Meteorological Department (TMD), Department of Water Resources (DWR) and Land Development Department (LDD) for providing information and data set.

Worapong Lohpaisankrit

---

[1] DAAD (in German: Deutscher Akademischer Austauschdienst)

[This page intentionally left blank]

Abstract

Flash flood hazard maps provide an essential support to mitigate flash flood risk. One of the main challenges in estimating flash floods is the scarcity of reliable runoff data. This is because flash floods are extreme phenomena, which can destroy monitoring equipment.

In the literature, various approaches have been proposed to estimate degrees of flash flood hazard. Most of them were developed in industrialized countries (e.g., the US and European countries) where large inventories of historic flash floods are available. However, none of them is proved to deliver a most suitable solution for flash flood hazard assessment in ungauged basins. Needless to say, it is more challenging to conduct flash flood studies in developing and emerging countries where detailed hydro-meteorological data are often limited.

Owing to unfavorable data conditions, it is difficult to prove which approach is more suitable than another one. This dissertation fills this gap by introducing a new spatial index-based approach to map potential flash flood hazard areas at high spatial resolutions preferably for ungauged basins but also for any basin, which needs a quick assessment of its flash flood potential.

For comparison purposes, another approach based on integrated hydrologic and hydrodynamic models was applied to model five sub-basins of the upper Nan River basin in Thailand, which suffered from flash floods in the past. Reliable meteorological and hydrometric data are missing in these sub-basins. Consequently, the entire upper Nan River catchment was modeled with the help of the hydrologic modeling system PANTA RHEI developed by the University of Braunschweig and a regionalization technique. This procedure allowed simulation of flash floods occurring in its sub-basins. Results of the PANTA RHEI model based on synthetic rainfall series were applied to drive a 2D hydrodynamic model in order to estimate spatial distributions of flash flood hazard in the five sub-basins. These results served as a reference for the development of the new approach.

For the new approach, physiographic variables of the five sub-basins were computed from available digital elevation models. These variables were linearly combined and automatically weighted by means of principal component analysis (PCA) to determine flash flood potential indices (FFPI). On the basis of the PCA, the first four principal components were related to the distance from stream, elevation, flow length, and drainage density. The FFPI values based on the five sub-basins could be described well by a generalized extreme value (GEV) probability distribution. The 98th percentile of the GEV probability distribution (so-called "regional" lowest classifier) was found to be suitable for identifying flash flood extents in the five sub-basins by comparison to results of the integrated modeling approach. Furthermore, one of the five sub-basins (the Nam Rim watershed) was tested by different spatially uniform rainfall distributions. As a result, flash flood hazard maps produced by the integrated modeling approach corresponding to the different rainfall distributions were compared with flash flood hazard maps. These maps were produced by adjusting classifiers for categorizing FFPI values into five levels of flash flood hazard.

The spatial index-based approach was also examined and applied to four sub-basins of another catchment (the upper Ping River basin) in similar climate. These four sub-basins were chosen to produce flash flood hazard maps without the help of a combined hydrologic and hydrodynamic model. To test the transferability of the regional lowest classifier being suitable for the five sub-basins of the upper Nan River basin, the

classifier was used to delineate flood boundaries of the four sub-basins of the upper Ping River basin. The areas inside of the boundaries were found to be smaller than historic inundated areas extracted from the Landsat-7 satellite imagery and the classifier was adapted.

On the basis of all investigations, the new spatial index-based approach is considered to be an effective procedure for flash flood hazard assessment in any basin including ungauged basins. Future work will focus on developing an analytical routing equation to transfer rainfall information (observed actual and forecasted rainfall) in order to extend the spatial index-based approach towards a (real-time) forecast.

Zusammenfassung

Sturzflut-Gefahrenkarten bieten eine wesentliche Unterstützung, um Schäden durch Sturzfluten zu verringern. Eine der größten Herausforderungen bei der Abschätzung der Ausmaße von Sturzfluten ist die oft geringe und unzuverlässige Datenlage. Extreme Sturzflutereignisse sind einerseits messtechnisch schwer erfassbar und können andererseits bestehende in situ Messgeräte beschädigen oder zerstören.

In der Literatur werden verschiedene Ansätze vorgeschlagen, um den Grad der Sturzflutgefahr abzuschätzen. Die meisten Ansätze wurden für eine Anwendung in Industrieländern entwickelt (z. B. USA oder Europa), wo die Datenlage bezüglich historischer Sturzflutereignisse weniger unzulänglich ist. Keine der recherchierten Ansätze wurde jedoch als eine geeignete Lösung für den Einsatz in nicht bepegelten Einzugsgebieten angesehen. Darüber hinaus ist die Aussagekraft solcher Ansätze in Entwicklungs- und Schwellenländer durch die oft sehr schwache Datenlage sehr begrenzt. Aufgrund dieser ungünstigen Datenbedingungen ist die Anwendung bestehender Ansätze für die Abschätzung der Sturzflutgefahr limitiert.

Die vorliegende Dissertation füllt diese Lücke durch die Einführung eines neu entwickelten, räumlich differenzierten, indexbasierten Ansatzes, um potentielle Gefahrengebiete mit hoher räumlicher Auflösung auch für unbepegelte Einzugsgebiete zu adressieren.

Zu Vergleichszwecken wurde ein weiterer, auf integrierten hydrologischen und hydrodynamischen Modellen basierender, Ansatz zur Modellierung von fünf Teileinzugsgebieten des oberen Nan-Flusseinzugsgebiets in Thailand verwendet. Diese Gebiete waren in der Vergangenheit von Sturzfluten betroffen. Zuverlässige meteorologische und hydrometrische Daten fehlen in diesen Teileinzugsgebieten. Durch Anwendung des von der Universität Braunschweig entwickelten hydrologischen Modellsystems PANTA RHEI und einer Methodik zur Regionalisierung von Eingangsdaten sowie Modellparametern wurde das gesamte obere Nan-Flusseinzugsgebiet zunächst modelliert. Die auf synthetischen Niederschlagsreihen basierenden und mit PANTA RHEI ermittelten Scheitelabflüsse wurden als Randbedingung für ein 2D hydrodynamisches Strömungsmodell verwendet. Mit diesem integrierten Modellansatz wurde die räumliche Verteilung der Sturzflutgefahr in den fünf Teileinzugsgebieten des oberen Nan bestimmt und diente als Referenz für den neuen Ansatz.

Grundlage des indexbasierten Ansatzes ist die Ermittlung von physiogeographischen Variablen aus digitalen Höhenmodellen. Diese Variablen wurden für die fünf Teileinzugsgebiete des oberen Nan linear kombiniert und automatisch mittels einer Hauptkomponentenanalyse (PCA) gewichtet, um räumlich differenzierte Indices des Sturzflut-Potenzials (FFPI) zu ermitteln. Auf Grundlage der Hauptkomponentenanalyse wurden die folgenden Variablen als die vier wichtigsten Hauptkomponenten identifiziert: Abstand zum Gewässer, topographische Höhe, Fließlänge sowie Gewässerdichte. Die FFPI-Werte aus den Teileinzugsgebieten konnten durch die allgemeine Extremwertverteilungsfunktion (GEV) gut beschrieben werden. Das 98% Perzentil dieser Extremwertverteilung (als “regional” niedrigster Klassifizierer) erwies sich durch Vergleich mit Ergebnissen des integrierten Modellierungsansatzes als gut geeignet, um die räumliche Ausdehnungen historischer Sturzfluten in den Testgebieten zu beschreiben. Darüber hinaus wurde eines der fünf Teileinzugsgebiete (Nam-Rim-Einzugsgebiet) mit räumlich unterschiedlichen Niederschlagsverteilungen belastet und die durch den integrierten Modellierungsansatz erzeugten Hochwassergefahrenkarten mit den entsprechenden Ergebnissen des indexbasierten Ansatzes im Detail verglichen. Die

indexbasierten Karten wurden durch Variation der Klassifizierung der FFPI-Werten ermittelt.

Zusätzlich wurde der indexbasierte Ansatz auf vier Gebiete des oberen Ping-Einzugsgebietes mit ähnlichen klimatischen Verhältnissen angewendet. Diese Einzugsgebiete wurden ausgewählt, um Sturzflut-Gefahrenkarten ohne die Hilfe eines mathematischen Modells (d.h. eine kombinierte hydrologische und hydrodynamische Modellierung) zu erzeugen. Um die Übertragbarkeit des regional niedrigsten Klassifizierers, der für die fünf Teileinzugsgebiete des oberen Nan-Flussgebiets geeignet ist, zu testen, wurde dieser zunächst auch für die Ping Einzugsgebiete verwendet. Die dadurch ermittelten Gefahrengebiete waren kleiner als die historischen Überschwemmungsgebiete, die aus Satellitenbildern des Landsat-7 extrahiert wurden und der Klassifizierer musste adaptiert werden.

Zusammenfassend lässt sich festhalten, dass auf der Grundlage aller Untersuchungen der neue, räumlich differenzierte und indexbasierte Ansatz ein effizientes und aussagekräftiges Verfahren für die Sturzflut-Gefährdungs-analyse in beliebigen, auch unbepegelten Einzugsgebieten ist. Zukünftige Arbeiten sollten sich auf die Entwicklung einer analytischen Routing-Gleichung zur Übertragung von Niederschlagsinformationen (be-obachtete aktuelle Daten und Vorhersage) konzentrieren, um den räumlichen indexbasierten Ansatz in Richtung einer Sturzflutvorhersage zu erweitern.

Contents

List of Figures

List of Tables

[This page intentionally left blank]

1 Introduction

1.1 Background

Floods occur when a body of water overflows and spreads across a piece of land that is usually dry. One of the most common floods is riverine floods. For research purposes, riverine floods are broadly referred to river floods and flash floods. River floods occur following a causative event, such as heavy rain and snowmelt, when watercourses do not have the capacity to accommodate the excess water. Consequently, the excess water overflows the banks of rivers and flows into low-lying floodplain areas adjacent to the rivers. In large rivers with relatively flat watersheds, water levels of main rivers tend to rise slowly because of an exceptional widespread and sustained rainfall. Their natural floodplains may remain inundated for several days, or even months (Tingsanchali and Karim, 2005).

Flash floods deviate from river floods in extent and duration. The term “flash” in context of hydrology denotes something that occurs quickly but it does not mean something that moves quickly (Kron, 2002). Therefore, flash floods are often characterized by an instantaneous increase in level and velocity of a flowing water body after the causative event. They are localized phenomena that normally occur in streams and small catchments with a drainage area covering a few up to several hundreds of square kilometers (Kelsch et al., 2001). Such small catchments, especially in mountainous terrains, respond quickly to intense storms because of steep slopes and impermeable surfaces. In order to differentiate between flash floods and river floods, a threshold in lag time[1] of approximately six hours is often used (Georgakakos, 1986; Sweeney, 1992; WMO, 1994; NOAA, 2010; Hapuarachchi et al., 2011).

Flash flooding is one of the most natural destructive disasters which can imperil many lives and cause losses of infrastructure and property. As mentioned above, that flash floods are usually provoked by very intense rainfall, they can also occur as results of dam failures. These flash floods due to dam breaks or sudden releases seem to be exceptional phenomena and are not considered in this dissertation. Instead, focus is given to flash floods triggered by intense rainfall.

1.2 Statement of problems

Rainfall-induced flash floods can happen almost anywhere. They often occur in tropical countries (Schumacher and Johnson, 2006). In tropical regions such as Thailand, destructive flash flood events have resulted in human casualties and destruction of houses and infrastructure in the past decades. These events are frequently associated with debris flows and landslides (Gruntfest and Handmer, 2001), especially in the northern parts of Thailand characterized by rugged topography and steep slopes. These rugged topography and steep slopes favor the quick onset of runoff when heavy rain falls on their surfaces. This quick onset

---

[1] Lag time is the time difference between the peak precipitation and and the peak flood

of runoff is a typical characteristic of flash floods that makes them difficult to forecast in limited warning times.

From a scientific point of view, the most important aspect of runoff generation is a combination of hydro-meteorological and hydrologic processes. Usually, available long-term hydro-meteorological observations and historic information may be used to predict extreme river floods. However, flash floods generally are caused by space-time distributed rainfall at small local scales. On account of their sudden development, conventional raingauge networks may not be sufficient to capture these events properly (Gaume et al., 2004). In addition, reliable runoff measurements at the sub-basin scale are usually inadequate. Consequently, existing knowledge on flash floods contains scientific gaps in the prediction accuracy.

In many developed countries such as the US and many European countries, gaps in scientific knowledge to improve the prediction accuracy of flash floods may be minor than the ones in developing and emerging countries such as Southeast Asian countries. This is mainly due to the difference in data availability. On the one hand, large inventories of flash flood events over Europe exist (Gaume et al., 2009). As a result, numerous flash flood studies to improve the understanding of these extreme events have been conducted in many European countries. Some of these studies have developed approaches for flash flood assessment in gauged catchments and have attempted to apply these approaches in ungauged catchments (e.g., Velasco-Forero et al., 2006; Borga et al., 2008; Rozalis et al., 2010; Garambois et al., 2015). However, there is still no clear view, which approach is preferable to assess the impact of flash floods in ungauged catchments even in European countries (Adamovic et al., 2016). On the other hand, in developing and emerging countries, there is very little availability or a lack of flash flood data (i.e., reliable runoff measurements). Thus, the development of an approach for flash flood assessment in these countries is a challenge. Moreover, if flash flood studies in ungauged catchments are based on the approaches developed from gauged catchments in different climate and geographic settings, reliability of their results may be questionable. Thus, there is an urgent need for developing an approach that mainly requires data (i.e., digital elevation maps) that are often available and that can be applied to every climate and geographic setting.

1.3 Current situation

In the current situation, flash flood investigations and descriptions still lack proficient methodology, mainly because these events are poorly measured and understood. One of the main issues is the scarcity of reliable runoff and high-resolution rainfall data in both spatial and temporal scales (Marchi et al., 2010). This issue precludes the use of traditional methods for characterizing flash floods (Ruiz-Villanueva et al., 2013).

Over the last few decades, various methods have been developed in order to understand flash floods in poorly gauged and ungauged basins including data-driven approaches and hydrologic modeling approaches. The data-driven approaches are often examined in basins where long-term recorded data of rainfall and runoff exist. It is then assumed that they will be applicable for ungauged basins as well (Artigue et al., 2012). The other approach is the hydrologic modeling approach. Physically based distributed hydrologic models have been acknowledged as important research tools to help understand mechanisms of flash floods and can be used to deal with the lack of quantitative information in ungauged basins as well. The advantage of the distributed hydrologic models is that their input parameters are often related to the physical attributes of the basins (Ogden et al., 2000; Foody et al., 2004). Thus, hydrologic models have been applied in flash flood studies by various researchers (e.g., Rulli and Rosso, 2002; Reed et al., 2007; Braud et al., 2010). Even though many advanced hydrologic models are available, the lack of temporal high resolution of rainfall deprives researchers of accurately estimating potential of flash floods.

The aforementioned approaches provide “rough” estimations (i.e., at basin scales) on flash flood hazard assessment, which may not be enough for flood risk assessment and management. Generally, it prefers to have flash flood hazard assessment at spatial scales. To identify flash flood hazard areas, common approaches can be categorized by two main aspects. The first is an integrated hydrologic and hydrodynamic modeling approach. The advantage of this approach is that flood zones at high resolutions can be delineated because of complex runoff routing algorithms in hydrodynamic models. As a result of computational expense, the hydrodynamic models are often applied to downstream areas of a basin or a short stream of a few kilometers (Kourgialas and Karatzas, 2014; Zhang et al., 2015; Lohpaisankrit et al., 2016). This approach may provide meaningful results if observed runoff and rainfall data are sufficiently available. Nonetheless, the accuracy obtained from applying the integrated modeling approach in small ungauged basins seems to be questionable.

The other approach develops flash flood potential (sometimes referred to as flash flood susceptibility) indices based on statistical analysis to delineate flash flood hazard zones. This index-based approach superimposes grid-based layers of physiographic and hydrologic variables in order to construct flash flood potential indices (FFPI). All variables are usually assumed to be equally important to predict flood-prone areas (Smith, 2003; Minea, 2013). For the reason that different physiographic and hydrologic variables have different influences on the flash flood occurrence, the method of Fernández and Lutz (2010) provides different weights to the variables based on “expert” opinions. This subjectivity may cause uncertainties in the results (Le Cozannet et al., 2013). In an attempt to minimize uncertainty in weighting, multivariate statistical analysis has been applied. For instance, Youssef et al. (2016) combined two statistical techniques of frequency ratio (FR) and logistic regression (LR) to predict flash flood susceptibility in space. Costache and Zaharia (2017) used FR and weights-of-evidence (WoE) to form FFPI for flash flood hazard mapping. In the two latter mentioned works, relative efficiency of their approaches was evaluated in comparison predicted areas prone to flash floods with the characteristics of known flooded and non-flooded locations. These approaches have some drawbacks to analyze classes of each flood-influencing variables before performing the weighting procedures. For example, the number of classes or a range of values in each class could be done according to expert knowledge.

In recent years, some research works (e.g., Kazakis et al., 2015; Khosravi et al., 2016) attempted to extend the aforementioned index-based approach by including rainfall variables. However, their analyzes seem to be flawed. For example, Khosravi et al. (2016) used three different bivariate statistical analysis methods – namely, Shannon’s entropy, statistical index, and normalization weighting approaches to produce flash flood susceptibility maps. These three approaches showed an excellent performance according to their data set of historic flood locations. From this literature, interesting questions concerning a weighting factor to rainfall variables have popped up because rainfall variables were given a lower weighting factor and some higher (sometimes highest) amount of rainfall was assigned a weight of zero. Thus, it is necessary to further investigate the influence of rainfall variables on an index-based approach, which is discussed later in this dissertation.

Up to the present, there is little published research on flash flood assessment in developing and emerging countries (e.g., Thailand, Vietnam and Malaysia). Van Ngo (2000) assessed flash floods in terms of quantity on time but not on space in Vietnam. In Tehrany et al. (2013)'s study, it was not flash flood assessment, but flood susceptibility assessment in Malaysia. Physiographic and rainfall variables were analyzed and weighted by using two different approaches: a combination of FR and LR models and a rule-based decision tree (DT) approach. The drawbacks of the FR and LR models have been discussed above. One of shortcomings in applying DT is that it involves expert knowledge.

As it has been described above, some approaches have few drawbacks in flash flood hazard mapping (or flash flood susceptibility mapping). In addition, some of them may lead to misunderstanding of flash floods for one who has little knowledge of the hydrology due to the analysis of rainfall variables. Thus, it is necessary to fill these research gaps by developing a concept for flash flood potential assessment.

1.4 Objectives and research approaches

This dissertation is aimed at contributing to a better understanding of flash flood potential assessment under poor data availability. To assess flash flood potential, a new spatial index-based approach is developed to explore the potential of deriving significant flash flood indices by integrating available spatially hydrologic and geomorphologic data sets, as well as integrated hydrologic and hydrodynamic modeling tools. The spatial index-based approach focused on assessment of both rural and urban flash flood potential in tropical regions, in particular in Thailand. In order to accomplish the above aim, it is necessary to fulfill the following research objectives, which are derived on the basis of a comprehensive literature analysis about present flash flood assessment in Chapter 2.

(i) Construction of a framework of integrated hydrologic and hydrodynamic models for flash flood potential assessment (FFPA):

  - to create synthetic flash flood events using synthetic rainfall events;
  - to build up integrated hydrologic and hydrodynamic models for simulating potential flash floods;
  - to produce flash flood hazard maps for people in order to assist risk management planning.

(ii) Development of a concept of FFPA based on flash flood potential indices (FFPI):

  - to formulate a robust strategy for flash flood hazard mapping based on available geomorphologic data sets and advanced statistical techniques;
  - to examine physiographic variables, which influence the occurrence of flash floods;
  - to compare results from FFPA with flash flood hazard maps produced by the integrated hydrologic and hydrodynamic models.

(iii) Examination of suitability of FFPA for the other regions:

  - to evaluate the concept of FFPA with observed flash floods of the regions.
  - to use the insight gained with the concept of FFPA to pinpoint implications for flood risk management.

The selected research approaches for FFPA consists of two principal phases. The first phase involves a traditional integrated hydrologic and hydrodynamic modeling approach. Besides the integrated modeling approach applied in the first phase, a new spatial index-based approach for the assessment of flash flood potential was developed by using the principal component analysis (PCA), a multivariate analysis technique. Both approaches were applied to the upper part of Nan River basin (see Chapter 4) in order to develop a concept of FFPA. An overview of the concept is shown in Figure 1.1.

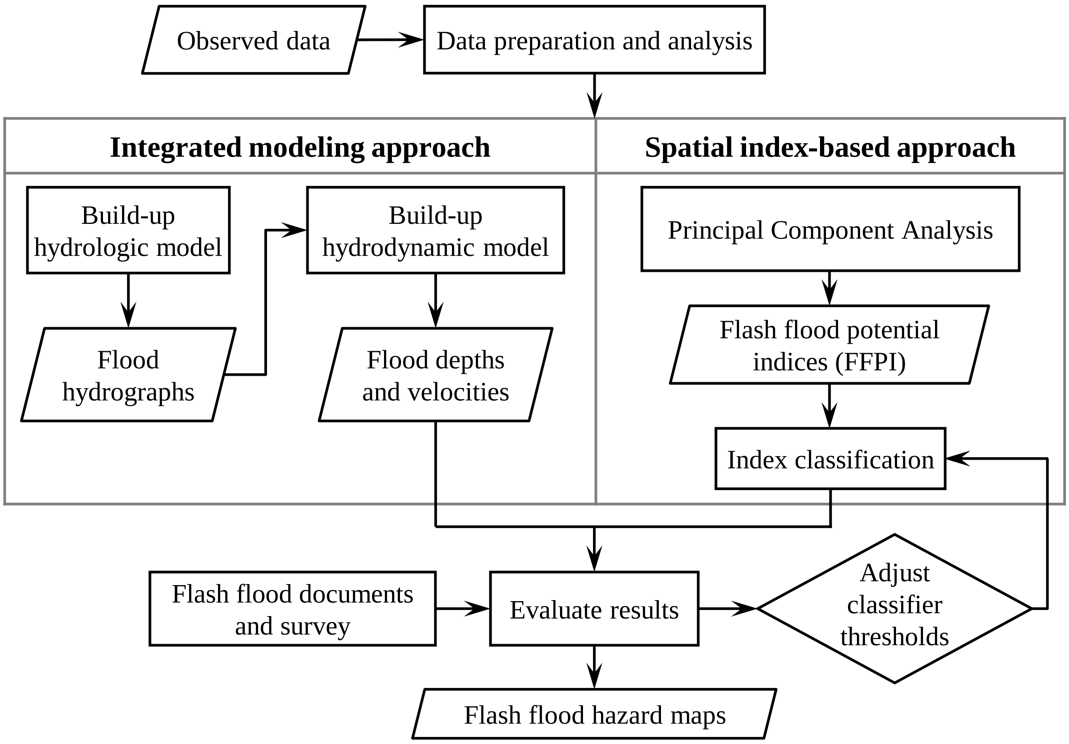


Figure 1.1: Overview of the concept for the assessment of flash flood potential: (left) integrated hydrologic and hydrodynamic modeling approach and (right) spatial index-based approach.

In the first phase, a hydrologic model was applied to the upper Nan River basin because there was no runoff observation in its sub-basins but observed discharge data at the main river were available. To simulate runoff in the sub-basins, the hydrologic model was used to transform rainfall data. Results (i.e., flood hydrographs, flood peaks) of the hydrologic model were provided as upstream boundary conditions of a hydrodynamic model. This hydrodynamic model was applied to five sub-basins of the upper Nan River basin in order to estimate flash flood areas, flood depths, and velocities. These outcomes from the hydrodynamic model were used to compute and produce flash flood hazard maps. One of advantages in the integrated hydrologic-hydrodynamic approach is that it can produce different flash flood hazard maps based on different extreme rainfall scenarios after it is calibrated and/or validated. Afterwards, the spatial index-based approach derived flash flood potential indices (FFPI) to determine areas prone to flash floods. To evaluate results of the spatial index-based approach, the values of FFPI were classified by classifier thresholds and then used to delineate flash flood hazard zones. These flash flood hazard zones were compared with different levels of flash flood hazard on maps produced by using the integrated modeling approach. Moreover, the concept of FFPA was developed.

In the second phase, only the spatial index-based approach was applied to the upper part of Ping River basin (see Chapter 4) in order to examine suitability of the concept of FFPA. The computation of FFPI for sub-basins of the upper Ping River basin was the same as in the first phase. However, this research work was limited by the lack of modeling results on flash flood hazard zones to compare. Therefore, areas prone to flash floods delineated by the spatial index-based approach were compared with inundated areas extracted from satellite imagery (see Section 6.5). To examine the the concept of FFPA, the classifier thresholds that were found to be suitable for the sub-basins of the upper Nan River basin were tested in sub-basins of the upper Ping River basin. The results are discussed in Chapter 7.

1.5 Scope and limitations of the study

The author seeks to explore flash flood potential assessment of sub-basins in the case of missing runoff observations. The sub-basins with drainage areas less than $350\,\mathrm{km}^2$ are considered because these sub-basins have a high potential of flash flood occurrence (Marchi et al., 2010). In addition, the main focus of this dissertation is to develop an approach to spatially assess flash flood potential for ungauged basins. Thus, this research work focuses to the use often available data (i.e., digital elevation maps) for developing the approach. This newly developed approach (i.e., a spatial index-based approach) based on statistical analysis delivers flash flood potential indices for determining the degree of flash flood hazard in space. As the dynamics of rain-driven flash floods are far beyond statistical analysis capabilities, flash flood potential indices are limited to a combination of “static” variables derived from digital elevation maps.

A major obstacle of this dissertation is that an integrated hydrologic-hydrodynamic modeling approach could not be proposed because this modeling approach requires extensive data (e.g., observed rainfall, runoff, high-flood marks) to calibrate and/or validate hydrologic-hydrodynamic models. However, this hydrologic-hydrodynamic modeling approach was applied in order to build up a scientific argument when only “inaccurate” data is available. As mentioned above, that there was no runoff record in the study sub-basins, a hydrologic model with a simple regionalization approach was applied to their parent catchment where discharge of its river was recorded. The underlying assumption of the regionalization approach is that the parent catchment and its sub-basins will behave identically, from a hydrologic point of view when they extend over a homogeneous climate zone. The parent catchment and its sub-basins were considered to be homogeneous in an extreme climate condition based on regional frequency analysis of annual maximum daily rainfall data. Moreover, calibrated model parameters of the parent catchment were not too sensitive. Therefore, the calibrated model parameters can be transferred to its sub-basins in order to estimate peak discharge of flash floods in these sub-basins. Moreover, owing to the lack of runoff data in the sub-basin, a hydrodynamic model is limited to a determination of boundary conditions based on outputs of the hydrologic model and assumed downstream levels. The calibration results of the hydrodynamic model were according to high-flood marks obtained from the author’s field investigations. It is assumed that there has been no land deposition and/or land subsidence between flooding events and the field investigations.

In addition to the result comparison between the developed index-based approach and the integrated modeling approach, the results of the index-based approach were compared to inundated areas obtained from satellite images. These inundated areas are assumed to display the flooding extents during the peak.

2 Literature Review

2.1 Benchmarks used in determining flash flood occurrence

2.1.1 Flooding flow

One conservative aspect defines a "flooding flow" (Q_p) as the bankfull flow (Q_{bf}). It is usually calculated by using the Manning-Strickler formula under two main assumptions: (i) a steady, uniform flow condition and (ii) a wide rectangular channel cross-section. The Manning-Strickler's flow resistance formula (Chow et al., 1988; Mosley and McKerchar, 1993) is expressed as follows:

$$Q_p = Q_{bf} = k_{St} B_{bf} D_{bf}^{5/3} S_f^{1/2} \tag{2.1}$$

where

Q_p	is the flooding flow [m^3 s^{-1}];
Q_{bf}	is the bankfull flow [m^3 s^{-1}];
k_{St}	is the Manning-Strickler friction coefficient [m$^{1/3}$ s^{-1}];
B_{bf}	is the channel top width at bankfull [m];
D_{bf}	is the hydraulic depth at bankfull [m];
S_f	is the friction slope approximated by the local slope of the channel [m/m].

To compute the bankfull discharge, it is obviously required to have information on channel cross-sectional parameters; however, this information is often not available for small streams. In unsurveyed streams, estimates of these parameters can be derived from other available digital elevation maps (DEMs) through regional regression relationships between the cross-sectional parameters and other basin and/or stream characteristics (Hapuarachchi et al., 2011). Simple power function relationships between channel hydraulic geometric characteristics of natural streams (e.g., channel width, mean flow depth, and velocity) and variation of discharge were first introduced by Leopold and Maddock (1953). The Leopold and Maddock (1953) approach was extended by Dunne and Leopold (1978) to develop regional curves that relate bankfull discharge and related bankfull channel geometry to catchment area for the upper Green River basin, Wyoming (the US). In general, the development of the regional curves requires observed values of downstream hydraulic geometry for a large number of rivers representing a wide variation of parameters. Although the concept of regional curves was applied across the entire United States, Bieger et al. (2015) justified that related bankfull channel geometric characteristics provide more reliable regional curves for estimating discharge than drainage area.

An alternative definition of flooding flow is on the basis of recurrence intervals for bankfull discharge. Research on bankfull discharge in the US has found that bankfull discharge occurs approximately within the range of one to two years for most streams (Dunne and Leopold, 1978; Henderson, 1966; Castro and

Jackson, 2001). To compute a flow of a certain recurrence interval, observed flow records with a long period of about 30 years or more are necessary. In the absence of observed flow data, rainfall-runoff models can use precipitation and catchment characteristics to simulate flow series (Reed et al., 2007). Carpenter et al. (1999) recommended a two-year recurrence interval flow (Q_2) as a useful indicator for flood warnings at ungauged sites because the Q_2 is often a little higher than bankfull discharge and can cause damage. To derive a regional curve for Q_2, the authors carried out a regression analysis on regional data such as drainage area, stream length, and average channel slope. They reported that the regional curve relating to drainage area and average channel slope for Q_2 had fairly strong correlation coefficient of 85% for their selected streams in Iowa. For streams in the southern California mountain region, Modrick and Georgakakos (2014) reported a range of values from 0.15 to 3.3 years, with an average value across this region of one year, produce bankfull discharge.

2.1.2 Runoff thresholds

Runoff thresholds (sometimes referred to as rainfall thresholds) have been used as main indicators of the flash flood occurrence and magnitude in many regions. In the United States (US), flash flood warnings have been provided by the National Weather Service (NWS) since the 1970s (Georgakakos et al., 1997). The NWS routinely depends on flash flood guidance (FFG) computations, and issues flash flood warnings based on the comparison of FFG with rainfall amounts. Generally, FFG refers to the depth of rainfall accumulated over a given duration, uniformly distributed over a considered basin that is capable of generating bankfull flow conditions at the basin outlet (Sweeney, 1992). When flows exceed bankfull flow at channel capacity, flooding starts. Thus, the bankfull flow has been used as a runoff threshold for deciding the flood occurrence.

The FFG approach is typically used to issue a warning for ungauged basins (Sene, 2013). Early stages of this FFG approach have developed estimates of runoff thresholds from digital terrain data and stream surveys. For example, a procedure for determining runoff thresholds proposed by Carpenter et al. (1999) was based on hydrologic and hydraulic principles. The authors assumed that basins linearly respond to rainfall excess. Therefore, runoff thresholds, which were used to indicate flooding on basin outlets, were assumed to be linearly related to rainfall excess for a specific duration. According to these assumptions, runoff thresholds can be equated by using linear unit hydrograph theory as follows:

$$R_{th} = \frac{Q_p}{q_{pR}A} \tag{2.2}$$

where

R_{th} is the runoff threshold [mm];
Q_p is the flooding flow [cms];
q_{pR} is the unit hydrograph peak for a specific duration t_R [cms mm^{-1} km^{-2}];
A is the basin area [km^2].

The flooding flow (Q_p) can be computed by using estimators such as statistical and physical quantities. An estimator is the two-year recurrence interval flow based on annual flood series. This flow is a reasonable approximation of bankfull flow (Carpenter et al., 1999). In ungauged basins, the estimation of bankfull flow can be achieved by using basin characteristics or channel geometry and roughness characteristics (Chow et al., 1988).

The applicability of the above-mentioned procedure developed by Carpenter et al. (1999) may be limited by the assumptions of the unit hydrograph theory (Martina et al., 2006). For instance, the linear rainfall-runoff transformations may not provide satisfactory predictions of the direct surface runoff for small basins because of a nonlinear relationship between rainfall excess and basin responds (Wang et al., 1981). However, a unit hydrograph approach under the assumption of uniform rainfall excess on the whole basin could be considered reasonable (Reed et al., 2002).

Apart from their extensive use in the United States, the concept of runoff thresholds has been applied and evaluated in many regions. For example, in Europe, research conducted as part of the FLOODsite[1] project aims at improving the ways to measure rainfall quantities from weather radar and providing guidelines based on the FFG method (Borga et al., 2008).

A rainfall depth (i.e., runoff threshold) that generates bankfull conditions at the outlet can be computed by a hydrologic model running in an inverse mode (Norbiato et al., 2008). For instance, Georgakakos (2006) run a hydrologic model with different rainfall and initial soil moisture conditions in order to determine surface runoff volume that can cause bankfull flow at the outlet of a given watershed. In addition, the simulated runoff volume, rainfall volume, and rainfall durations were expressed in analytical relationships that can be used to provide FFG estimates.

2.2 Synthetic rainfall generation

2.2.1 Disaggregation of rainfall

Rainfall data at a fine temporal resolution (sub-daily) are necessary for various hydrologic applications. However, observed sub-daily rainfall data are scarce in many regions. Where sub-daily rainfall data exist, the available periods of the data are generally too short for water resources projects such as flood forecasting and urban drainage design (Gaume et al., 2007). On the other hand, long records of daily rainfall are normally available. Therefore, many studies attempt to address the lack of the fine temporal resolution data by disaggregation of daily rainfall data. The main concept of the temporal rainfall disaggregation is that statistical properties of historic coarser temporal resolution (e.g., daily) rainfall are transferred to synthesize finer temporal resolution (e.g., hourly) rainfall.

Over the last decades rainfall disaggregation models have been developed. In a review of rainfall disaggregation models, Sharma and Mehrotra (2010) categorized commonly used models for rainfall disaggregation into three groups: statistical models, Poisson cluster process based models, and scale invariance models. The first group of rainfall disaggregation models is based on statistical analysis. These models use the concept of dimensionless mass curves to characterize storm intensity patterns (Huff, 1967). For example, dimensionless mass curves of variables (e.g., the number of rainfall event per day, rainfall depth, duration, and starting time) at finer time scales are fitted theoretical probability distribution functions. Afterward, these functions are transferred to temporally disaggregate rainfall data at coarser time scales (Connolly et al., 1998). In addition to the use of theoretical probability functions for rainfall disaggregation, Koutsoyiannis and Foufoula-Georgiou (1993) developed empirical probability functions to fit dimensionless mass curves of recorded rainfall depth and duration in Charala (Greece). Molnar and

---

[1] FLOODsite is an interdisciplinary project supported by the European Commission.

Burlando (2005) stated that the above-mentioned variables may not enough to preserve the characteristics of extreme rainfall events. They recommended to consider the sequence of pulses and intermittency in rainfall.

Rainfall disaggregation models in the second group are based on point-process stochastic models. Two types of the point-process stochastic models – namely, the Bartlett-Lewis and Neyman-Scott models, have been described by Rodríguez-Iturbe et al. (1987) and have been applied worldwide in the context of rainfall simulation. These models investigate the intermittency of the temporal rainfall process and simulate rainfall by a Poisson cluster process. Rainstorms arise according to a Poisson process and each storm is characterized by a random number of cells. Each cell is characterized by a random intensity and a random duration (Bo et al., 1994; Wheater et al., 2005). The advantages of the point-process stochastic models are that rainstorms can be summarized by a few parameters and can be resampled at any desired aggregation scale (Koutsoyiannis and Foufoula-Georgiou, 1993). By contrast, using the point-process stochastic models to reproduce extreme values seems to be deficit when return periods are greater than the length of the data set (Onof et al., 2000).

In the last group, scale invariance models convert rainfall time series at a coarser resolution into a finer resolution through multiplicative cascade processes. The analysis of these processes generally consists of two stages: an examination of the moment scaling function and a selection of an appropriate probability distribution for the weights (Sharma and Mehrotra, 2010). Molnar and Burlando (2005) used cascade-based models for the disaggregation of 1280-min into 10-min rainfall data for a recoding period of 20 years in Zurich (Switzerland). Their models performed well in preserving rainfall intermittency, but not the variability in simulated rainfall extremes. Similarly, Pui et al. (2012) obtained a poor performance of a cascade model for rainfall disaggregation. Their cascade model overestimated extreme rainfall when it was used to downscale daily rainfall to hourly rainfall. In contrast, Güntner et al. (2001)'s temporal rainfall disaggregation model based on the principles of random multiplicative cascade processes could reproduce extreme rainfall in a semi-arid climate region. They report that transferability of model parameters in time appear to be associated with larger uncertainty in the region because of the higher interannual variability of rainfall and lower percentage of rainy intervals.

It can be noted that the authors cited here developed rainfall disaggregation models in areas where fine temporal rainfall data exist for long periods. So that, these rainfall disaggregation models can be evaluated by statistical comparison between finer temporal resolution (e.g., hourly) and coarser temporal resolution (e.g., daily) rainfall data. This may serve as a drawback because their applicability is limited to areas where at least some recorded rainfall data at fine time scales are available.

2.2.2 Design hyetographs

A design hyetograph is a synthetic rainfall temporal distribution that is usually determined by analysis of observed rainstorm events. However, observed fine time scale (sub-daily) rainfall data are usually inadequate in many developing and emerging countries. Thus, many flash flood studies (e.g., Kim and Choi, 2011; Lohpaisankrit et al., 2016), especially in these countries, have often adopted design hyetographs developed from the others. To develop design hyetographs from recorded rainfall data, various methods have been proposed. According to Veneziano and Villani (1999), most of these methods can be classified into four categories:

(i) Specification of simple geometric shapes relative to a single point of the intensity-duration-frequency (IDF) curves or depth-duration-frequency (DDF);

(ii) Usage of patterns in accordance with the entire IDF curves;

(iii) Normalized profiles obtained directly from rainfall records;

(iv) Simulation from stochastic rainfall models.

In the first category, approaches for designing hyetographs are based on a single point of the IDF curves or DDF curves. A classic approach uses a rectangular (uniform) shape in combination with an average rainfall intensity obtained from the IDF curves. This approach is often used with the Rational method (as described in Section 3.2.2) for flood protection design (e.g., Pilgrim and Cordery, 1993). This approach has been noticed that the total precipitation volume is underestimated when the rainfall intensity of a rectangular shape is derived from the IDF curves (Veneziano and Villani, 1999). An alternative geometric form is a triangular shape (Chow et al., 1988).

Approaches in the second category use patterns relative to the entire IDF curves. In other words, design hyetographs are developed from IDF curves. Examples of approaches in this category are the Chicago approach (Keifer and Chu, 1957) and the alternating block method (Chow et al., 1988).

In the third category, approaches derive dimensionless hyetographs by normalized rainfall profiles (e.g., Huff curves (Huff, 1967) and SCS rainfall temporal distributions (USDA, 1986)). The dimensionless hyetographs represent time distributions of storm rainfall. For instance, Huff (1967) developed dimensionless hyetographs on the basis of heavy storms in a 12-year period on areas ranging up to about 1000 km^2 in Illinois (US). According to the approaches in this category, large sample sizes are required in order to produce stable average rainfall profiles. Moreover, it appears that the dimensionless hyetographs depend on the temporal rainfall patterns in the region considered (Veneziano and Villani, 1999). The concept of the dimensionless hyetographs has been applied in flash flood studies. For example, Kim and Choi (2011) used dimensionless hyetographs based on Huff curves in combination with local IDF curves as input into a lumped hydrologic model for flash flood prediction in a small ungauged watershed of Korea.

In the last category, approaches generate hyetographs based on stochastic rainfall models. In general, the stochastic rainfall models derive parameters related to statistical properties of examined rainfall data. For example, Koutsoyiannis and Foufoula-Georgiou (1993) estimated the parameters by using a stochastic process of storm arrivals (e.g., a Poisson cluster process as discussed above in Section 2.2.1) and a set of distribution functions fitted to rainfall durations and depths for generating design hyetographs.

Of the aforementioned four categories of the design hyetograph methods, the methods in the first two categories establish the design hyetograph by taking into account characteristics of both the rainfall process and the basin response. In doing so, these two methods use the IDF curves or DDF curves as a synthetic representation of the rainfall climate and determine the rainfall duration related to the response time of the basin or to peak runoff rate. On the contrary, the methods in the last two categories define the design hyetograph based solely on the rainfall process (Veneziano and Villani, 1999).

2.3 Existing approaches for estimating flash flood runoff

The estimation of flash flood runoff is a challenging process because detailed hydro-meteorological data are normally limited in poorly gauged or ungauged basins. Sometimes hydrologic characteristics of flash floods are partially known in poorly gauged basins. For the reason that stage gauges may be non-functional during extreme flood events, the extreme characteristics cannot be completely obtained. Moreover, the number of rain gauges in poorly gauged basins is not sufficient to fully capture storm events at the appropriate spatial and temporal scales (Norbiato et al., 2007; Koutroulis and Tsanis, 2010). Therefore, numerous approaches have been developed in order to study and estimate extreme floods in the poorly gauged and ungauged basins.

2.3.1 Data-driven approaches

Data-driven approaches – namely, transfer functions and artificial neural networks (ANNs), usually use available time-series data to characterize the model system. Thus, the reliability of the approaches relies on input (e.g., rainfall) and output (e.g., runoff) measurements (Toukourou et al., 2011). One aspect that these approaches have been widely used is that they do not require any prior knowledge about hydrologic conditions of a watershed such as initial soil water content (Wagener et al., 2013).

Transfer functions

Transfer functions are primarily based on observations and can be formed by using an empirical approach. An early example for estimating peak discharge is the Rational method (Mulvaney, 1851; See Section 3.2.2). This Rational method estimates peak discharge from three main variables: catchment area, a maximum average rainfall intensity, and an empirical coefficient C. The empirical coefficient C is often used as a constant parameter. However, Beven (2012) argued that the empirical coefficient C is not a constant parameter and is required to take account of the nonlinear relationship between antecedent conditions and the rainfall distribution and the resulting runoff generation. In Koutroulis and Tsanis (2010)'s study, an nonlinear empirical equation was developed on the basis of regression analysis. This nonlinear empirical equation was defined as a function of the total precipitation, its duration, and its standard deviation to estimate flash flood peak discharge in the Giofiros basin (Greece) with a drainage area of 186 km^2. They recommended that the nonlinear empirical equation can be applied to poorly gauged basins because coefficients of the equation for an individual basin have to be specifically computed from its characteristics. Generally, empirical approaches based on regression analysis are developed for gauged watersheds to estimate discharge from characteristics of the watersheds (Rosbjerg et al., 2013). These characteristics can be assessed from geographic information systems and meteorological statistics (Mosley and McKerchar, 1993). The underlying assumption of these approaches is that watersheds are considered to be similar with respect to climate and their physical characteristics. Under this assumption, watersheds will behave similarly in a hydrologic sense (Wagener et al., 2013).

Another classic approach develops transfer functions based on unit hydrograph theory (Dooge, 1959). It has been demonstrated that hydrologic response of a basin is controlled by its geomorphologic characteristics (Horton, 1945; Beven and Kirkby, 1979; Marchi et al., 2010). Therefore, the links between basin response to extreme rainfall and geomorphologic variables have been considered by a number of published studies (e.g., Wood and Hebson, 1986; Diakakis, 2011). One of the most widely used methods to discover the rainfall-runoff relationships is the geomorphologic unit hydrograph (GUH) derivation (Rodríguez-Iturbe and Valdés, 1979; Wood and Hebson, 1986; Jain and Sinha, 2003). Diakakis (2011) suggested that this GUH method is valuable for flood hazard assessment, particularly in ungauged basins, because the geomorphologic parameters of a basin can be derived from a digital elevation model that is available worldwide. Ghoneim et al. (2002) preferred, for estimating flash flood peak discharge in an ungauged basin, the Snyder (1938)'s synthetic unit hydrograph. Owing to the lack of runoff observations, the rainfall was assumed to uniformly spatially distribute over its sub-basins. Consequently, peak discharge was dependent solely on the characteristics of the sub-basins.

Artificial neural networks (ANNs)

Artificial neural networks (ANNs) categorized as data-driven approaches are ones of widely-used paradigms for flash flood forecasting. These approaches do not account for physical processes (Toukourou et al., 2011). Over the past years, research has been carried out on a variety of artificial neural networks (ANNs) techniques for flood forecasting because of their specific properties. The first property of ANNs is that a complex nonlinear relationship can be established from a set of known variables through training processes (Schmitz and Cullmann, 2008). The second property is that ANNs do not require priori assumptions for modeling built-up (Artigue et al., 2012).

For purposes of flash flood forecasting, Piotrowski et al. (2006) performed three ANNs techniques – namely, nearest neighbors, multi-layer perceptron neural networks (MLPNs), radial-basis function neural networks (RBFNs) to examine rainfall-runoff relationships based on gauged rainfall and river discharge data of the Nysa Klodzka catchment (Poland). They found the nearest neighbors and RBFNs approaches performed better than the MLPNs approach for the training data. To prevent a poor generalization capability of an ANNs model, (Dreyfus, 2005) suggested to use a regularization approach such as an early stopping and a cross-validation along with the ANNs model.

Currently, ANNs based on the multi-layer perceptron neural networks (MLPNs) with a a regularization approach have been widely used in streamflow and flash flood estimations. For example, Coulibaly et al. (2000) applied a MLPNs model with an early stopping criterion for regularizing the neural networks to predict streamflow in long-term time series. Their MLPNs model with the regularization approach can provide better prediction accuracy than the MLPNs model without regularization performed. Ghumman et al. (2011) also recognized that performing ANNs with a regulation approach to model the rainfall-runoff relation can improve the performance of the ANNs. They regularized their ANNs model by using a cross-validation. The performances of the ANNs model were assessed to be better or at least equivalent to a conceptual rainfall-runoff model. In Toukourou et al. (2011), 15 flash flood events occurred in the Gardon d' Anduze catchment (France) were used to construct a ANNs model that consisted of two main layers: recurrent neural networks and feedforward neural networks with time delays. The reported results showed satisfactory in flash flood simulations. Toukourou et al. (2011)'s work on some improvements of ANNs for flash flood forecasting was complemented by Artigue et al. (2012)'s study of ANNs sensitivity analysis to rainfall. Sensitivity analysis of a ANNs model based on the MLPNs to rainfall inputs was performed by using observed 17-year rainfall and discharge data from three raingauges and discharge gauges on the Mialet basin (a sub-basin of the Gardon d' Anduze catchment) with a drainage area of 220 km^2. They demonstrated that their ANNs model was low sensitive to rainfall.

In general, the main advantage of ANNs for estimating flash flood runoff is that ANNs models can provide reliable forecasted runoff without initial information (e.g., initial soil moisture condition and initial flow) once the ANNs models were trained and constructed (Schmitz and Cullmann, 2008; Artigue et al., 2012). On the other hand, the main disadvantage of using the data-driven approaches based on ANNs for estimating flash flood runoff is that they require long-term recorded data of rainfall and streamflow to train and validate. Thus, the data-driven approaches may be difficult to apply in small poorly gauged catchments because of short-term recorded data or no data to build the rainfall-runoff relation (Hapuarachchi et al., 2011). Moreover, these data-driven approaches may lead to large uncertainty in predictions when condition (e.g., climate and land use) changes because hydrologic response and processes are not taken into consideration.

2.3.2 Hydrologic modeling approaches

Typically, flash floods caused by intense rainfall events are nonlinear phenomena. These phenomena can be characterized by using physically based models, which aggregate the surface-water components of the hydrologic processes (Beven, 2001). A spatially distributed hydrologic model has been underlined to be the most useful tool for predicting flash floods in the absence of detailed field measurements because its input parameters are related to physical attributes of basins (Ogden et al., 2000; Foody et al., 2004) and it uses a hydrologic similarity concept (Bonnifait et al., 2009).

A number of authors have integrated conceptual and process-based hydrologic methods to predict flash floods and improve understanding of these events. For example, Foody et al. (2004) applied a semi-distributed hydrologic model with the SCS concept to predict flash floods in ungauged watersheds. Main parameters of their model were derived from topographic maps, field survey, and land cover maps. These parameters were not calibrated. Similarly, Yucel (2015) used an uncalibrated model based on the SCS method (as described later in Section 3.2.2) to estimate flash flood hydrographs in the Ayamama basin (Turkey) from different precipitation data sets because there was no records of streamflow measurements. Moreover, Rozalis et al. (2010) used an uncalibrated model based on the SCS concept and the kinematic wave model to predict flash floods in a small Mediterranean watershed. They found that the model preformed well for extreme events, while it overestimated for low and moderate events. Conversely, Gaume et al. (2004) argued that it is necessary to calibrate the value of the curve number (CN), especially in lumped hydrologic models, owing to the spatial heterogeneity of discharges in a watershed. Therefore, CN values tend to vary with respect to rainfall and soil moisture conditions (Riedel, 2004). Another application of integrated hydrologic modeling approaches is in the Zhang et al. (2015)'s study. They preferred, for flash flood simulations in a midsize Chinese watershed, a calibrated hydrologic model with the GUH method. They suggested to GUH method for a basin with limited rainfall and runoff data because its parameters can be derived from geomorphologic data of the basin.

For the purpose of flash flood forecast, Estupina-Borrell et al. (2006) developed the distributed hydrologic model MARINE[2]. There was no calibration performed; however, model parameters were derived from a priori information and expert knowledge concerning geospatial data (e.g., topography, soil, and land cover) and measurements (e.g., hydraulic roughness coefficient and cross section of the river). The MARINE model without parameter calibration was tested by Braud et al. (2010) on a set of small ungauged

---

[2] MARINE stands for Model of Anticipation of Runoff and INondations for Extreme events

basins ranging from a few km^2 to about $100\,\mathrm{km}^2$ in France. To evaluate the model, post-flood field estimates of peak discharges and timing were compared with the model simulations based on one extreme flood event. The performance of the model was satisfactory but not for all the tested basins. It is noted that a large impact on the simulated hydrologic response may be due to soil characteristics and initial water content conditions. In a gauged watershed, Roux et al. (2011) calibrated a few parameters of the MARINE model and used Monte-Carlo simulations to perform sensitivity analysis of these parameters. They advised that the model's parameterization for flash flood forecasting, especially in ungauged basins, can improve by applying regionalization methods. In keeping with the Roux et al. (2011)'s suggestion on parameter regionalization, Garambois et al. (2015) applied two regionalization methods – namely, a geographic proximity approach and a physical similarity approach, to transfer calibrated parameters from gauged to ungauged basins. Their study was based on a large and various flash flood data set (117 flash flood events) of the French Mediterranean region. They revealed that the physical similarity approach produce better results than the geographic proximity approach for the flash flood data set.

For parameter regionalization, Velasco-Forero et al. (2006) derived multiple linear regressions from basin physical properties and model parameters. The parameters of a distributed conceptual hydrologic model were calibrated independently for each sub-basins of the Besòs catchment (Spain) and were correlated with physical properties of all sub-basins on the basis of the multiple linear regressions. Afterwards, the relationships developed in the Besòs catchment were used to estimate model parameters in sub-basins of the Anoia catchment, which is approximately 50 km far from the Besòs catchment. The estimated model parameters failed to produce hydrographs that were in a good agreement with observed hydrographs. They concluded that their methodology of parameter regionalization is not suitable for catchments that have different physical characteristics. Likewise, Garambois et al. (2015) held the view that the hydrologic response is not always consistent between neighboring catchments.

In an attempt to regionalize parameters, Adamovic et al. (2016) analyzed parameters of a semi-distributed hydrologic model for small basins of the Ardeche basin (France) using factor analysis and hierarchical classification. Their results pointed out that geology is a predominant factor of hydrologic variability. Therefore, they estimated three geology recession parameters from gauged basins and transferred them to ungauged ones in order to simulate flash floods. The model results showed improvements in some of the ungauged basins and suggested further analysis such as sensitivity analysis of simulation results according to parameter regionalization.

To prove an efficiency of a spatially distributed hydrologic model in ungauged basins, Bonnifait et al. (2009) applied TOPMODEL[3] to reconstruct a catastrophic flood event in the Gardon watershed (France). Parameters of the model were calibrated with the stream gauge, which drains an area of $1090\,\mathrm{km}^2$. Furthermore, the hydrologic parameters upstream were assumed to be homogeneous and the calibrated parameters were adopted to simulate flash flood hydrographs in small sub-catchments upstream of the stream gauge. According to the simulations in these upstream sub-catchments, distributed hydrologic modeling with a simple parameterization, which transfers calibrated parameters of a catchment to its sub-catchments, seems efficient in reproducing the spatial hydrologic response.

---

[3] TOPMODEL is a TOPography based hydrological MODEL developed by Beven and Kirkby (1979)

It can be observed that the above-mentioned approaches were established in industrialized countries (e.g., US and European countries) where large inventories of flash flood events exist (Gaume et al., 2009). As a result, numerous research on flash floods in these countries is found in the literature. On the contrary, there is very little published research (e.g., Van Ngo, 2000) on flash floods in emerging and developing countries (e.g., Thailand and Vietnam) because of poor data availability. This situation may not be suitable to perform regionalization analysis because of too few study cases. Supporting this view, Garambois et al. (2015) wrote that it is difficult to find donor catchments that contain enough relevant hydrologic information to reproduce particular catchment behavior even in a large data set. This emphasizes a large number of case studies and runoff observations required to investigate and improve regionalization performances in predicting runoff in ungauged catchments.

2.3.3 Post-flood survey

When discharge data directly measured with a current-meter are missing, post-flood survey can provide useful information for indirectly estimating peak discharge (Gaume et al., 2004). A number of so-called “indirect” hydraulic methods have been developed such as slope-conveyance, slope-area, step-backwater, and critical-depth methods (Webb and Jarrett, 2002). These methods rely upon the empirical Manning’s equation (sometimes referred to as the Manning-Strickler equation) first proposed by Philippe Gauckler in 1867.

In general, the indirect hydraulic methods require high-flood marks and stream cross-sections in order to estimate flood peak discharges. The most commonly used method for indirect estimates in ungauged basins is the slope-conveyance method, introduced by Scott and Gravlee (1968). This is because the slope-conveyance method enables a large number of peak discharge values for a single flood event to be evaluated (Gaume and Borga, 2008). The slope-conveyance method is usually applied under two principal assumptions (Marchi et al., 2009):

- Uniform flow, where the channel bed, the water-surface profile, and the energy-grade lines are parallel;
- One cross-section represented for all the other possible cross-sections in the channel reach.

Gaume and Borga (2008) expressed that the uniform flow assumption may provide fair velocity estimates for relatively high-gradient channels in the upstream basins if the selected cross-section is far enough (some hundred meters) upstream and downstream from specific conditions that can cause backwater effects. According to this slope-conveyance method, peak velocity can be estimated by means of the one-dimensional Manning-Strickler equation (as described in Equation 3.10 in Section 3.2.1). This equation requires one cross-sectional wetted area with the flood marks to compute values of peak discharge. For this reason, the slope-conveyance method has been widely used for discharge estimation in ungauged basins (e.g., Marchi et al., 2009; Amponsah et al., 2016). However, the slope-conveyance method with one surveyed cross-section may induce large errors in discharge estimates because cross-sections along the river channel vary as a result of erosion and/or channel aggradation (Webb and Jarrett, 2002). Hence, Chow (1959) noted that the quality of peak discharge estimation based on slope-conveyance method relies on the selection of the cross-sections and of the Manning-Strickler roughness coefficients.

As flow is usually nonuniform in natural channels, the slope-conveyance method is expected to produce large errors in discharge estimates (Webb and Jarrett, 2002). To cope with the uniform flow, Costa and Jarrett (2008) preferred the slope-area method developed by Dalrymple and Benson (1968) to estimate peak discharges of the 21 extraordinary flood events in the United States. This slope-area method is based on a combination of the Manning-Strickler, the conservation of mass, and the conservation of energy equations. This combination requires at least two cross-sections for computing velocity heads at upstream and downstream locations, and can provide more than one discharge.

Another method based on the same combination of the hydraulic equations as the slope-area method is the step-backwater method (Chow, 1959). However, the slope-area method computes the discharge from known water-surface elevations, while the step-backwater method computes water stage from discharge. Step-backwater analysis can be performed by using computer programs such as HEC-2 and HEC-RAS (Webb and Jarrett, 2002). For example, Ruiz-Villanueva et al. (2013) applied the HEC-RAS program to estimate peak discharges in the Arroyo Cabrera catchment (Spain). There was no measurement of streamflow. To estimate possible peak discharges, the results of the step-backwater method were compared with ones of an indirect discharge estimation based on the slope-conveyance method.

Post-flood field data (i.e., peak discharge and time of peak) are valuable indicators, especially for small ungauged basins, to evaluate hydrologic models for flash flood simulation. However, these data are not sufficient for a full assessment of the simulated hydrologic response in terms of flow and soil saturation dynamics (Braud et al., 2010). Amponsah et al. (2016) added that the indirect estimates of peak discharges based on the post-flood field data may not be accurate because of geomorphic changes. The cross-sectional geometry surveyed is different from the one during the peak flood since erosion or deposition processes may occur after the peak flood. Moreover, uncertainty in the indirect estimates may be due to a selection of roughness coefficient, especially for flash floods (Lumbroso and Gaume, 2012). Values of the roughness coefficient (Manning-Strickler roughness coefficient) can generally be taken between 5 and 100 depending on the river cross-section characteristics (Gaume and Borga, 2008). Thus, it is important to keep the possible uncertainty in indirect estimates of peak discharges in mind and apply a range of methods to support the estimates.

Several studies of post-flood survey have used dendrogeomorphic techniques to support estimation of peak discharges in ungauged basins. Ballesteros-Cánovas et al. (2011) designed a methodology that combined dendrogeomorphic techniques with a hydraulic model to reconstruct peak discharge of flash floods in an ungauged basin. The dendrogeomorphic techniques estimate peak discharge by using paleostage indicators (i.e., scars on trees) on the basis of the step-backwater method. These paleostage indicators were statistically compared with water depths simulated by the hydraulic model for estimating their corresponding peak discharges at sites.

2.4 Flash flood hazard assessment

Generally, the hazard of flooding is defined by its probability of occurrence and its magnitude (Kron, 2002). Therefore, flash flood hazard assessment requires integrated understanding of meteorology, hydrology, geomorphology (Mani et al., 2014). Spatial results of the flash flood hazard assessment are usually presented by using indicators on the map that represent the degree of hazard such as extreme, significant, moderate, low, and very low (e.g., Tingsanchali and Karim, 2005; Mani et al., 2014; Kazakis et al., 2015; Costache and Zaharia, 2017).

A fundamental element of flash flood hazard assessment is the identification of areas susceptible to flooding. Recently, two main approaches commonly used to spatially assess flash flood hazard are: (i) coupled hydrologic-hydrodynamic methods and (ii) index-based approach.

2.4.1 Coupled hydrologic-hydrodynamic methods

Estimation of flash floods may be challenging because they often occur in the areas where reliable runoff data are limited (Gaume et al., 2009). As a result of the scarcity of these data, many studies have used various approaches to estimate flood peaks such as data-driven approaches and hydrologic modeling approaches (see Section 2.3). These approaches can be categorized as hydrologic methods. On the one hand, the hydrologic methods are useful for flash flood estimation. In ungauged catchments (i.e., lack of runoff observations), the hydrologic methods can be used to simulate flash floods of various exceedance probabilities based on rainfall data normally available. On the other hand, the hydrologic methods are often inadequate for comprehensive analysis of inundation areas (Smith and Ward, 1998) because they follow a one-dimensional procedure, while the morphology of a river and its floodplain are unstable. Therefore, recent approaches account for hydrologic methods coupled with hydrodynamic methods to simulate the flood extent (e.g., Koutroulis and Tsanis, 2010; Kourgialas and Karatzas, 2014; Zhang et al., 2015).

One of the main advantages in applying hydrodynamic models is that continuous flood boundaries can be simulated. Generally, these hydrodynamic models are essential to be calibrated relative to specific historic flood events. Once the hydrodynamic models are calibrated, they can be used to estimate flood extents and inundations corresponding to different flooding scenarios. For the identification of areas susceptible to flooding, hydrodynamic modeling methods are usually applied on the basis of three underlying assumptions (Domeneghetti et al., 2013):

- Hydrodynamic (i.e., hydraulic) models are capable of reproducing the hydraulic behavior of the river and its floodplains;
- Model parameters (e.g., roughness coefficients) calibrated for a specific event are considered suitable for a range of flooding scenarios;
- Upstream and downstream boundary conditions (e.g., hydrographs and rating curves) are "error-free".

Coupled hydrologic and hydrodynamic methods have been extensively used in flash flood hazard assessment. In particular applications, hydrologic forcing (i.e., rainfall) is transformed by a hydrologic method into flood peaks or flow hydrographs of given return periods. These hydrologic outputs are often used as an upstream boundary condition for a hydrodynamic model (Grimaldi et al., 2013). In hydrodynamic flow simulation, estimated floods of various exceedance probabilities obtained from hydrologic methods are transformed into more useful information such as depth of flooding, velocity of floodwater, and duration of flooding for flood hazard assessment (Tingsanchali and Karim, 2005). For flash flood hazard assessment, depth of flooding and velocity of floodwater are mainly considered (Penning-Rowsell et al., 2005; Jonkman and Penning-Rowsell, 2008; Xia et al., 2011). For instance, Kourgialas and Karatzas (2014) used a calibrated hydrologic model to simulate flash flood hydrographs at an upstream boundary of a hydrodynamic modeling area (i.e., downstream area of the Koiliaris River basin, Greece). These hydrographs were routed through the river and its floodplains by using a calibrated MIKE11 (quasi-2D) hydrodynamic model. As a result, flood depth and velocity were obtained and then were used to compute flash flood hazard rates by using an empirical equation (Ramsbottom et al., 2003; Jonkman and Penning-Rowsell, 2008). Zhang et al. (2015) estimated peak discharge by using a calibrated hydrologic model and computed depth of flooding by using the Manning's equation. The computed flood depth was solely the parameter used for flash flood hazard mapping along the Xiapu River (China) and its floodplains.

Although hydrodynamic modeling methods have proved to be helpful in providing useful information for flood hazard assessment, results of these methods may not be reliable because of four dominant sources of uncertainty: input data (e.g., boundary conditions), calibration data (e.g., high-flood marks and flood extent maps), model parameters (e.g., roughness coefficients), and model structural uncertainty (Hunter et al., 2005; Di Baldassarre, 2012).

To assess the flood hazard degree for people, previous studies (e.g., Abt et al., 1989; Karvonen et al., 2000) conducted experiments to examine related-flood hazard factors – namely, the water velocity and depth, which caused human instability in floodwater. Some studies adopted experimental results to develop empirical or semi-quantitative equations to estimate level of hazard to people because of flooding. Ramsbottom et al. (2003) proposed an empirical formula as a function of flood depth, flow velocity, and debris potential. This empirical formula had been developed on the basis of historic flood events in the UK and their impacts on people to support flood hazard mapping. The original empirical formula of Ramsbottom et al. (2003) was modified by Ramsbottom et al. (2006) according to the mechanics-based experiments of Abt et al. (1989) and Karvonen et al. (2000) on human instability in flood flows. The revised formula (i.e., Equation 6.4 as shown in Section 6.4) presented by Ramsbottom et al. (2006) generally appears to be an accurate criterion for assessing and mapping hazard to people because of flooding (Kvočka et al., 2016). However, Cox et al. (2010) argued that the Ramsbottom et al. (2006)'s empirical formula suffers from three serious weaknesses:

- The empirical formula was derived from the average values of flow regime[4] obtained from the experimental data sets. It is not necessarily valid with respect to the general population because most people may experience instability at their first exposure to a flood hazard.
- The debris factor (DF) was not supported by experimental testing.
- The empirical formula has no upper depth limit, and thus large depths at low velocities of flood flows are not necessarily considered as hazardous. This is impractical because of these flow conditions in which a subject becomes buoyant, and therefore people are inherently unstable and safety may depend on swimming ability.

2.4.2 Index-based approach

Many studies have focused on using morphometric variables to improve the understanding of flash floods. The hydrologic response of a basin is controlled by its morphometric variables (Patton and Baker, 1976; El-Magd et al., 2010). In various studies, these morphometric variables have been used to estimate flood peaks

---

[4] flow regime is indicated as the product number of depth of flooding multiplied by the flood velocity

(Ghoneim et al., 2002; Masoud, 2004; Dawod et al., 2011; Zemzami et al., 2013) and to determine corresponding hazard levels of flash floods (Ghoneim et al., 2002; Youssef et al., 2011). The latter procedure has often been applied to map flash flood hazards in ungauged regions. Although the scarcity of reliable hydro-meteorological data limits flash flood studies, morphometric and physiographic features can often be derived from digital elevation models. Afterwards, these features can be used as indicators for mapping flash flood hazards.

In the scientific literature, there are two major groups of index-based approaches for flash flood hazard assessment. The first group assesses the degree of flash flood hazards at basin scales. However, the second group can identify flash flood hazard areas at higher resolutions (e.g., a 30 m×30 m grid). These two groups of index-based approaches are discussed in more detail in the following sections.

2.4.2.1 Flash flood hazard assessment at basin scales

There are a large number of published studies (e.g., Patton and Baker, 1976; Ghoneim et al., 2002) that describe the link between morphometric characteristics of basins and flooding impact. Ghoneim et al. (2002) identified flash flood hazards in the El-Alam basin (Egypt) by using relationships between morphometric variables and its hydrologic response. The hydrologic response (i.e., peak discharge) of the basin was estimated by using a synthetic unit hydrograph approach because the basin was ungauged. They found that the hydrologic response was highly related to seven morphometric variables – namely, basin area, drainage density, basin ruggedness, basin relief, relief ratio, maximum surface slope, and surface slope.

Numerous studies have attempt to analyze characteristics of basins and used them to assess flash flood hazards and/or potential. For example, Omran et al. (2011) computed five morphometric variables of the Dahab watershed (Egypt) as shown in Table 2.1 from a digital elevation model (DEM). In order to assess flood hazards for sub-basins of the watershed, values of the morphometric variables were classified by using normalization weights. Another example in Bajabaa et al. (2014) suggested that nine morphometric variables (basin area, drainage density, stream frequency, shape factor, slope, relief ratio, ruggedness ratio, texture ratio, and bifurcation ratio) have influences on the occurrence of flash floods. Some more applications and more examples of flash flood hazard assessment based on morphometric variables are presented in Table 2.1.

The approach used by the above-mentioned studies delineates flash flood hazard zones at basin scales. To clarify, the approach assigns one hazard degree for the whole basin. The flaws of this approach have been clearly recognized. In the Abdelkareem (2017)'s study, the approach was applied to determine degrees of flash flood hazard for sub-basins of the Asyuti watershed (Egypt). As shown above in Table 2.1, the twelve morphometric variables of the sub-basins were selected and their values were weighed by using linear equations in order to identify the degree of flash flood hazard to these sub-basins. The flash flood hazard zones (i.e., sub-basins) were compared with flash flood extents obtained from remote sensing data (i.e., satellite images). He reached the conclusion that the approach provides very rough estimation on flash flood hazard assessment because many areas in the sub-basins identified by a very high degree of flash flood hazard were not inundated.

Table 2.1: Selected literature on flash flood hazard assessment at basin scales

Morphometric variables	Weighting methods	Study area	Reference
Dd, F_{str}, *Rb*, R_{cir}, and R_{elong}	Normalization weights	Dahab basin, Egypt	Omran et al. (2011)
A, C_{com}, *Dd*, F_{shape}, F_{str}, H_b, *HI*, L_o, N_{geo}, N_{rug}, N_{strord}, R_{cir}, R_{elong}, S_m, S_r, T_{ns}, and T_{sl}	Linear equations	Feiran basin, Egypt	Youssef et al. (2011)
A, *Dd*, F_{shape}, F_{str}, I_{slp}, N_{rug}, *Rb*, R_{tex}, and S_r	Linear equations	Al Lith basin, Saudi Arabia	Bajabaa et al. (2014)
Dd, F_{form}, F_{str}, H_b, L_o, N_{rug}, *Rb*, R_{cir}, R_{elong}, R_f, R_{tex}, and S_r	Linear equations	Asyuti basin, Egypt	Abdelkareem (2017)

Note: basin area (A), compactness coefficient (C_{com}), drainage density (*Dd*), form factor (F_{form}), basin shape factor (F_{shape}), stream frequency (F_{str}), basin relief (H_b), hypsometric integral (*HI*), basin slope index (I_{slp}), length of overland flow (L_o), geometry number (N_{geo}), ruggedness number (N_{rug}), stream order number (N_{strord}), bifurcation ratio (R_b), circulation ratio (R_{cir}), elongation ratio (R_{elong}), infiltration ratio (R_f), texture ratio (R_{tex}), basin slope (S_m), relief ratio (S_r), total number of streams (T_{ns}), and total stream length (T_{sl}).

2.4.2.2 Flash flood hazard assessment at grid scales

At a high-resolution delineation, flash flood prone areas are identified by using the Flash Flood Potential Index (FFPI) initiated by Smith (2003) for flash flood monitoring and prediction in the Colorado River basin. Gridded physiographic data (e.g., see in Table 2.2) over the domain of interest were assigned with values. The highest value indicates the maximum flash flood potential. Afterwards, the physiographic features were equally weighted, and the integrated grid outcomes of FFPI were created. Finally, the FFPI values were classified with equal intervals into different flood hazard levels. The Smith (2003)'s approach on the determination of FFPI has been carried out by several authors like Zaharia et al. (2012) and Minea (2013). One of the shortcomings of this FFPI-based approach was that it did not take into account the temporal variability of soil moisture conditions, which is one of the most dominant factors that influences flash flood occurrence (Costache and Zaharia, 2017). Another shortcoming of the approach was that physiographic variables were equally weighted. This may not be the case that physiographic variables equally contribute to flash flood occurrence (Wahid et al., 2016).

Instead of giving equal weight to variables, Wahid et al. (2016) weighted physiographic variables unequally according to their importance in flash flood occurrence. In addition, statistical techniques such as analytic hierarchy process (AHP), frequency ratio, and weights-of-evidence, have been widely used to analyze and weight variables in order to map flood hazard. Fernández and Lutz (2010) performed a multi-criteria decision analysis based on AHP to map flood hazard in the two Argentinian cities of San Miguel de Tucumán and Yerba Buena. The major limitation of AHP is rather subjective as it relies on expert opinions, which is the main source of uncertainty (Le Cozannet et al., 2013). In an attempt to weight variables without involving expert opinions, Costache and Zaharia (2017) applied an approach combining two statistical techniques – namely, frequency ratio (FR) and weights-of-evidence (WoE), to construct FFPI for flash flood hazard assessment in the Basca Chiojdului River basin (Romania). These two statistical techniques were used to weight variables according to past flood locations. The weak point of this approach was that values of variables need to be classified into certain classes before using the weighting techniques. This classification procedure was based on expert knowledge.

Table 2.2: Selected literature on flash flood hazard assessment at grid scales

Physiographic variables	Weighting methods	Study area	Reference
Land use, soils, surface slope, and vegetation	Equally weights	Colorado River basin, US	Smith (2003)
Depth to groundwater table, distance from stream, land use, surface slope, and surface elevations	Analytic hierarchy process	San Miguel de Tucumán and Yerba Buena cities, Argentina	Fernández and Lutz (2010)
Land use, profile curvature, soils, and surface slope	Equally weights	Basca River basin, Romania	Minea (2013)
Land use, soils, surface slope, and vegetation	Unequally weights	Yunnan Provice, China	Zeng et al. (2015)
Land use, shape factor, soils, and surface slope	Unequally weights	Qaa Plain, Egypt	Wahid et al. (2016)
Aspect, convergence index, drainage density, L-S factor, land use, lithology, profile curvature, soils, and surface slope	Weights-of-evidence and frequency ratio	Basca Chiojdului River basin, Romania	Costache and Zaharia (2017)

As known that flash floods are mainly driven by rainfall, several studies attempted to assess flash flood potential (sometimes referred to as flash flood susceptibility) and/or flash flood hazard based on combined physiographic and rainfall variables. Kazakis et al. (2015) applied AHP to weigh physiographic and rainfall variables and then calculated a flood hazard index. Khosravi et al. (2016) used three different bivariate statistical analysis methods: Shannon's entropy, statistical index, and normalization weighting approaches to produce flash flood susceptibility maps of the Haraz watershed (Iran). These maps produced on the basis of ten variables – namely, distance from stream, geology, land use, normalized difference vegetation index, profile curvature, stream power index, surface elevation, surface slope, topographic wetness index, and rainfall. These variables were analyzed and weighted by the three statistical methods in order to compute indices for flash flood susceptibility mapping. The results were evaluated on the basis of historic flood points obtained from field surveys and documentary sources and proved capabilities of the three statistical methods for identifying flash flood locations. Kourgialas and Karatzas (2011) provided flood hazard maps of the Koiliaris River basin (Greece) based on six variables: flow accumulation, geology, land use, surface elevation, surface slope, and rainfall intensity. Values of the variables were classified into levels of flood hazard by using the natural breaks classification method and were weighted by expert knowledge. Their results showed that areas identified by very high flood hazard were consistent with historic flooded points. However, it is interesting to note that most authors who used index-based approaches to assess flash flood potential gave a low weighting factor to rainfall.

2.5 Conclusions from literature review

This section summarizes what one can learn from this chapter of literature review. There are four main parts: benchmarks used in determining flash flood occurrence, synthetic rainfall generation, existing approaches for estimating flash flood runoff, and flash flood hazard assessment.

Benchmarks used in determining flash flood occurrence

Bankfull discharge (Q_{bf}) is commonly used to decide the flood occurrence in gauged and ungauged rivers. In gauged rivers, bankfull discharge can be described in terms of its recurrence intervals. However, it is impracticable in ungauged rivers. Therefore, many studies have attempted to estimate bankfull discharge from rainfall observations through different methods such as empirical approaches (e.g., Chow et al., 1988; Carpenter et al., 1999) and hydrologic modeling approaches (e.g., Georgakakos, 2006; Norbiato et al., 2008).

Lessons learned and implications for this research work:

- The estimation of bankfull discharge requires information on channel cross-sectional geomorphology and a selection of roughness coefficients. Uncertainty in the results can be due to the selection of channel cross-sections and roughness coefficients.
- Bankfull discharge of bankfull stage is commonly used as an indicator of the flood occurrence at the outlet of a given basin.
- In this research work, bankfull stage was used as a downstream boundary condition for a hydrodynamic model (see Section 6.2.2).

Synthetic rainfall generation

In poorly gauged or ungauged basins, synthetic rainfall data are necessary in order to improve accuracy of the flood discharge estimation. Generally, rainfall observations at high temporal scales like hourly are not available in emerging and developing countries. Therefore, numerous studies (e.g., Connolly et al., 1998; Güntner et al., 2001; Wheater et al., 2005) with respect to disaggregation of daily rainfall have been conducted in where finer temporal resolution rainfall data exist and several rainfall disaggregation approaches were proposed. Some approaches based on statistical or stochastic models were successful in producing extreme rainfall characteristics at finer time scales (e.g., hourly) from coarser temporal resolution (e.g., daily) rainfall data for a specific region, but not for another region (Güntner et al., 2001). These approaches may not applicable to where high temporal scales of rainfall data are scarce for validating the results of daily rainfall disaggregation.

Lessons learned and implications for this research work:

- To construct a rainfall disaggregation model, long time series of both daily and sub-daily rainfall observations are needed.
- Once a rainfall disaggregation model based on statistical or stochastic techniques is constructed, it may be only valid in a region where it is developed.
- In this research work, observed daily rainfall time series were disaggregated into hourly rainfall time series by using temporal distributions of hourly rainfall observed at nearest neighboring and/or observed by a radar weather station (see Section 5.2).

Existing approaches for estimating flash flood runoff

Over the years, many approaches (e.g., data-driven techniques, hydrologic modeling techniques) have been proposed in the literature and aim at predicting flash flood runoff in ungauged basins. Most of them were developed on the basis of available rainfall and runoff observations. The data-driven techniques have one of advantages that data-driven models do not need any prior knowledge about hydrologic conditions (e.g., initial soil moisture and initial water content) after these models are calibrated and/or validated (Wagener et al., 2013). On the other hand, the models may be difficult to be constructed in areas where lack in data to form the rainfall-runoff relation (Hapuarachchi et al., 2011). In addition to data-driven techniques, spatially distributed hydrologic models have proved to be the most useful tool for predicting flash flood runoff in ungauged basins because their input parameters are related to physical characteristics of the basins (Foody et al., 2004). However, this is not always the case, as uncertainty in hydrologic modeling may be due to unknown soil moisture and water content conditions (Braud et al., 2010; Lumbroso and Gaume, 2012).

In an attempt to predict flash flood runoff in ungauged basins, some of the above-mentioned approaches have combined regionalization concepts. The underlying assumption of these concepts is that basins similar in climatology and physiography will behave similarly in a hydrologic sense (Wagener et al., 2013). Thus, a rainfall-runoff relation or model developed from one area could be used to predict runoff in another area if both areas have a hydrologic similarity. Moreover, transferring parameters of a hydrologic model between the both areas will be acceptable.

The application of the hydrologic models with regionalization approaches seems to provide some improvement in flash flood estimation but it is performed at a high cost basis (e.g., Velasco-Forero et al., 2006; Garambois et al., 2015).

Apart from estimating runoff by using the rainfall-runoff approaches, indirect hydraulic methods for peak flood estimation have been widely used, especially in ungauged basins (Gaume et al., 2004). These methods estimate peak flood discharges from high-flood marks and stream cross-sections through the empirical Manning-Strickler equation. The quality of peak discharge estimation depends upon the selection of the cross-sections and of the Manning-Strickler roughness coefficients (Chow, 1959). However, uncertainty in the indirect hydraulic methods for peak flood estimation is due to changes in cross-sectional geomorphology and a selection of the roughness coefficients (Lumbroso and Gaume, 2012; Amponsah et al., 2016). Thus, it is necessary to apply alternative methods to valid results of the indirect flow estimation.

Lessons learned and implications for this research work:

- Data-driven models can provide reliable runoff without initial hydrologic conditions (e.g., initial soil moisture and initial water content) but solely under the same conditions as they are constructed.
- Data-driven approaches may be difficult to apply in poorly gauged basins due to short-term recorded data or no data.
- Indirect hydraulic methods are interesting approaches for estimating peak discharge at the outlet of an ungauged basin. However, large uncertainty of these methods can be due to the selection of cross-sections and roughness coefficients.
- Hydrologic models may be the most practical tool for predicting flash floods in poorly gauged basins because model parameters are related to physical characteristics of the basins, which can be derived from available digital elevation maps.

- Hydrologic models usually need to be calibrated with runoff measurements. In ungauged basins, the application of hydrologic models with regionalization techniques can improve the reliability of the results.
- In this research work, a spatially distributed hydrologic model with a regionalization approach was used to estimate flash flood runoff at interior points of the study basin (see Section 6.2). This was due to flash floods occurred in sub-basins of the study basin where observed runoff data were not available.

Flash flood hazard assessment

Generally, the hazard of flooding is usually defined in terms of its probability of occurrence and its magnitude (Kron, 2002). In gauged streams with a long discharge record, flash flood hazard mapping can be done by using only hydrodynamic methods. On the other hand, in ungauged streams, probability of flash flood occurrence and flood peak discharge can be estimated from available rainfall data by using hydrologic methods. Afterwards, the outcomes of the hydrologic method are used as input to hydrodynamic methods in order to produce flash flood hazard maps.

Approaches for flash flood hazard assessment in ungauged basins can be categorized into two groups. The first group is coupled hydrologic-hydrodynamic methods. Using the first part of the coupled hydrologic-hydrodynamic methods (i.e., hydrologic methods), flash flood peak discharges at the outlet of a given basin can be estimated. The estimation of peak discharge provides "rough" information on flash flood hazard assessment. That is to say, it can be used to determine degree of flash flood hazard at catchment scales with respect to estimated peak discharge. However, estimated peak discharge is often used as the upstream boundary condition for the last part of the coupled hydrologic-hydrodynamic methods (i.e., hydrodynamic methods). In general, outcomes of hydrodynamic methods are such as depth of flooding, velocity of floodwater, and duration of flooding. The two former outcomes (i.e., flood depth and velocity) are usually used for flash flood hazard assessment (Jonkman and Vrijling, 2008; Xia et al., 2011). The major shortcoming of the methods in this first group is the collection and management of large amounts of data, an expensive and time-consuming process. Another shortcoming is that uncertainty in hydrologic and hydrodynamic methods is still issues in the assessment of flash floods in ungauged catchments (Lumbroso and Gaume, 2012; Di Baldassarre, 2012).

The last group of approaches for flash flood hazard assessment is index-based approaches. These index-based approaches can be classified into two main categories according to the spatial scale being considered. In the first category, index-based approaches assess degree of flash flood hazard at catchment scales (e.g., Bajabaa et al., 2014; Abdelkareem, 2017). These index-based approaches provide a rough assessment of flash flood hazard, which is comparable to that given by the hydrologic methods, as previously mentioned. However, index-based approaches do not evaluate flash flood hazard of a given catchment based on quantity of runoff at the outlet, but they evaluate flash flood hazard based on morphometric variables of a given catchment that have a high potential for flash flood occurrence. Index-based approaches in the other category provide higher spatial resolutions (i.e., grid scales) of flash flood hazard mapping (e.g., Smith, 2003; Costache and Zaharia, 2017). This flash flood hazard mapping on a grid basis is similar to that performed by the aforementioned hydrodynamic methods. The flash flood hazard mapping based the index-based approaches has been carried by numerous researchers. Most of them develop flash flood potential indices (FFPI), sometimes referred as to flash flood susceptibility indices, from physiographic variables over the areas of interest without considering rainfall influence. Other researchers (e.g., Kourgialas and Karatzas, 2011; Kazakis et al., 2015) attempted to develop the indices based on physiographic and rainfall variables. To assess the degree of flash flood hazard, the indices need to be weighted and classified. In doing so, statistical techniques (e.g., analytic hierarchy process, frequency ratio, and weights-of-evidence) have often used. It has been observed that the researchers who used index-based approaches based on the statistical techniques assigned a low weigh factor to rainfall. Therefore, it is necessary to investigate whether statistic analysis can cope with the dynamics of rain-driven flash floods and start an argument about whether rainfall can be statistically analyzed together with other "static" physiographic variables.

Lessons learned and implications for this research work:

- A coupled hydrologic-hydrodynamic method may be the most useful tool for flash flood hazard assessment in the areas where the lack of reliable runoff measurements is because runoff can be estimated by a hydrologic approach from rainfall.
- However, uncertainty of the coupled hydrologic-hydrodynamic method in ungauged basins may be due to four main sources: (i) runoff estimation based on a hydrologic approach (for boundary conditions of a hydrodynamic model), (ii) calibration data (e.g., flood discharges, high-flood marks and flood extent maps), (iii) model parameters (e.g., roughness coefficients), and (iv) model structures (Di Baldassarre, 2012).
- In this research work, a coupled hydrologic-hydrodynamic model was applied to map flash flood hazard for the comparison purposes (see Section 6.4).
- For flash flood hazard assessment in ungauged basins, alternative approaches are index-based approaches that determine the degree of flash flood hazard based on a combination of "static" physiographic variables and statistical techniques. These approaches are useful because their variables can be computed from a digital elevation map that is often available.
- Several researchers (e.g., Kourgialas and Karatzas, 2011; Tehrany et al., 2013; Kazakis et al., 2015; Khosravi et al., 2016) attempted to develop indices based on physiographic and rainfall variables for assessing the degree of flood hazard. The author raised a question why rainfall variables were given lower weights.
- In this research work, a new spatial index-based approach was developed to form flash flood potential indices (FFPI) based on a combination of static physiographic variables and a statistical technique. In addition, sensitivity analysis of the spatial index-based approach to spatial rainfall distributions was performed in order to investigate the influence of the rainfall distributions on the spatial index-based approach (see Section 6.3).

3 Theoretical Background

3.1 Definition of flash floods

Several definitions for flash floods are given by various scholars. Some of them are provided as follows:

- A flash flood is any flood that occurs at a particular area within a few hours after the causative event, which is either rainfall or sudden operational release of water from a reservoir. The time interval of **12 hours** between the time of the causative event occurrence and the time of the flash flood occurrence is adopted to distinguish flash floods from other river floods (Georgakakos, 1986).
- A flash floods is a flood that arises within less than **six hours** of heavy rainfall (Sweeney, 1992).
- Flash floods are results of rapidly responding small catchments, which have time of concentration shorter than **six hours**, to intense rainfall (WMO, 1994).
- A flash flood is a phenomenon, in which the hydrologic processes rapidly develop on the same spatial and temporal scales as the intense precipitation. This phenomenon causes the sudden increase in flood level and velocity (Kelsch et al., 2001).
- Sudden floods with little lead time are generally at a small local scale and often occur in unexpected places, which make warnings problematic (Gruntfest and Handmer, 2001).
- Flash floods are localized phenomena occurring in basins of few hundred square kilometers or less, with maximum response times of a few hours (Borga et al., 2007).
- A flash flood is a rapid and extreme flow of high water into a certain area, which begins within **six hours** of the causative event, such as intense rainfall, dam failure and ice jam. However, the actual time threshold may vary in different parts of a region (NOAA, 2010).
- Flash floods are extreme floods, which are generated by intense precipitation over rapidly responding catchments. The short responding times of the catchments within a few hours can be used to distinguish flash floods from regular floods (Brauer et al., 2011).
- In general, rainfall-induced flash floods are characterized by their rapid onset within **six hours** of rainfall, which limits opportunity to effectively cope (Hapuarachchi et al., 2011).

According to the above-mentioned definitions of a flash flood, a threshold of approximately six hours, which is the time of the delay between the rainfall event and the flood peak, often is used to differentiate a flash flood from a “normal” slow-rising flood. The time of the delay, which is the response time of a watershed to rainfall, depends on the hydrologic characteristics–such as terrain slope, land use, soil types and soil moisture–and the hydraulic processes of the watershed (NRC, 2005).

3.2 Hydrologic modeling

3.2.1 Catchment-scale hydrologic processes

Hydrology focuses on the distribution and circulation of water through various earth systems above and below the ground. This recurrent circulation, which is called the hydrologic cycle, can be splitted into two main phases: the atmospheric and the Earth's surface phases. The surface phase is the main interest in surface hydrology, which mainly concentrates on the interface linking meteorological and hydrologic processes.

For modeling rainfall-runoff processes, very simple approaches such as black-box, regression and empirical methods simplify the processes to a few parameters. For example, the Rational method proposed by Mulvaney (1851) determines peak runoff according to a runoff coefficient, a rainfall intensity, and a basin drainage area. On the other hand, more sophisticated hydrologic approaches normally account three chief rainfall-runoff processes: runoff formation, runoff concentration, and runoff routing (Baumgartner and Liebscher, 1996).

3.2.1.1 Runoff formation

The hydrologic mechanisms that are usually involved in runoff formation on a slope surface are shown in Figure 3.1. This runoff formation is associated with precipitation (falling of water on the ground surface). The water enters the hydrologic processes and exits a cross section of a given watershed as total runoff. The total runoff in streams can be generally divided into two components: direct runoff and baseflow. The direct runoff is the summation of overland flow and interflow. Overland flow occurs when the amount of rainfall exceeds the capacity of the soil to infiltrate water. Interflow is infiltrated water, which moves horizontally in the unsaturated zone. These overland flow and interflow processes deliver water to the stream within a few hours or about a day. The other part of the infiltrated water flows vertically downward to the saturated zone and eventually reaches the stream over longer periods such as weeks, months or even years. This part is known as baseflow (Tarboton, 2003).

Generally, surface runoff of a given watershed is defined as the difference between total precipitation and hydrologic abstractions. These hydrologic abstractions are processes for reducing total precipitation to effective precipitation. The processes by which precipitation is abstracted by the watershed mainly consist of interception, evaporation, infiltration, and surface or depression storage.

Interception

Interception is the process by which precipitation is captured by vegetation. The interception loss describes the proportion of precipitation that is routed as direct throughfall or is stored on vegetative surfaces and depressions, and is subsequently evaporated from them. The amount of Interception depends on rainstorm character, vegetation and climate conditions (Dunne and Leopold, 1978).

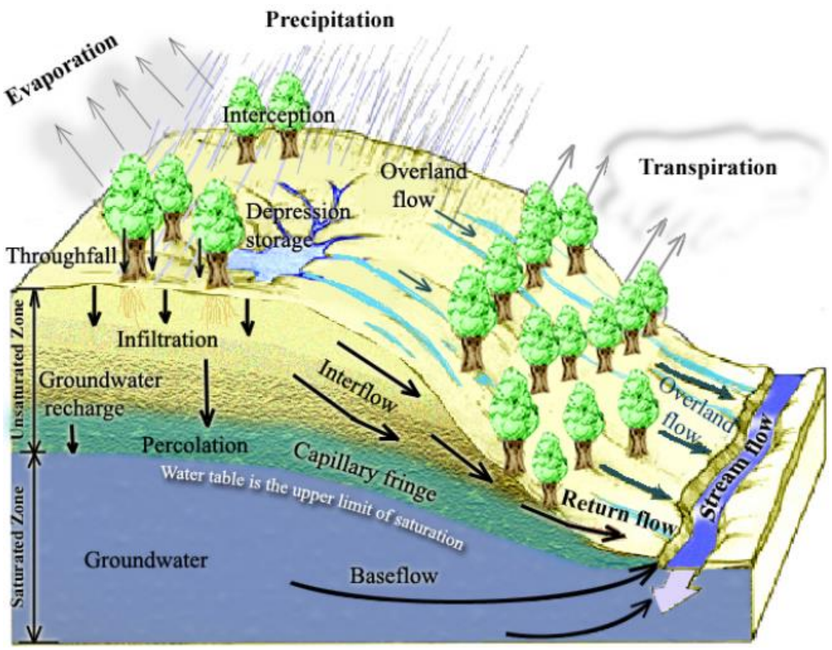


Figure 3.1: Hydrologic processes involved in runoff formation (taken from Tarboton (2003), modified).

Evaporation

Evaporation from open water surfaces is influenced by two main factors: the energy supply to provide the latent heat of vaporization and the ability to transport the vapor away from the water surface. The energy supply is mainly obtained from solar energy. The ability to transport the vapor away from the water surfaces is governed by the specific humidity gradient in the air above the surface and the wind speed over it.

Evaporation can be estimated by using an energy balance equation as shown in Equation 3.1, which is the combination method introduced by Penman (1948) that combined an energy budget with an aerodynamic equation. The former part of the Equation 3.1 represents the energy term and the latter part represents the aerodynamic term. This combination method for estimating evaporation is suitable for application to small areas with detailed meteorological data including net radiation, air temperature, air pressure, humidity, and wind speed (Chow et al., 1988).

$$E = \frac{A_s}{A_s + \gamma} E_r + \frac{\gamma}{A_s + \gamma} E_a \tag{3.1}$$

where

E	is evaporation rate [e.g., mm d^{-1}];
A_s	is slope of the saturated vapor pressure vs. temperature curve at mean air temperature [Pa K^{-1}];
γ	is psychrometric constant [Pa K^{-1}].
E_r	is evaporation rate based on net radiation [mm d^{-1}];
E_a	is aerodynamic evaporation [mm d^{-1}];

For estimating evaporation of watersheds with large areas or a scarcity of vegetation data, Priestley and Taylor (1972) simplified the Penman combination method by approximating the aerodynamic term in proportion to the energy term in Equation 3.1. Thus, this equation can be rewritten as the Priestley-Taylor equation:

$$E = \alpha \frac{A_s}{A_s + \gamma} E_r \tag{3.2}$$

where α equals to 1.3 according to Priestley and Taylor (1972). However, the value of α varies slightly from one site to another (Chow et al., 1988).

In addition to evaporation from open water surfaces, evapotranspiration is the process which consists of evaporation from soil and vegetation surface, and transpiration from vegetation. The evaporation directly occurs from soil and vegetation surface. The other is the transpiration whereby soil water is absorbed through plant roots, transported upwards through the plant and distributed into the atmosphere through stomata on the leaves.

Evapotranspiration is controlled by energy supply and vapor transport, which are the same factors governing evaporation from open water surfaces. Therefore, estimates of evapotranspiration rates are similar to the aforementioned methods for open water evaporation, with adjustments that account for vegetation and soil conditions (Monteith, 1973; Thompson, 1999).

Infiltration

The Horton infiltration model, which is a widely known empirical equation, can estimate infiltration capacity of the study area. The model was derived on the assumption that infiltration capacity is not usually constant and is reduced at the beginning of rain by effects from the energy of falling rain (Horton, 1945). It was concluded that infiltration begins at some rate f_0 and exponentially decreases to a steady state infiltration capacity f_c. This exponential decay of infiltration capacity is given by:

$$f_h = f_c + (f_0 - f_c)\, e^{-\beta t} \quad \text{for } 0 \geq t \geq t_c \tag{3.3}$$

where

$f_h\,(t)$	is the infiltration capacity at any time t [mm h^{-1}];
f_c	is the final state infiltration capacity occurring at $t = t_c$ [mm h^{-1}];
f_0	is the maximum infiltration rate at the beginning of a storm event ($t = 0$) [mm h^{-1}];
β	is the Horton's decay parameter, which controls the rate of decrease in the infiltration capacity [h^{-1}];
t	is time [h].

The cumulative infiltration F after time t for the Horton model is expressed as

$$F(t) = f_c t + \frac{(f_0 - f_c)}{\beta}\left(1 - e^{-\beta t}\right) \tag{3.4}$$

where the cumulative infiltration F is in millimeters.

3.2.1.2 Runoff concentration

Runoff concentration refers to the concentration of the rainfall excess, which is transferred to the watercourse. It is also used to describe the temporal displacement of individual flows in different soil storage reservoirs. The runoff concentration can be simulated on the basis of reservoir-routing methods. A reservoir-routing method with uncontrolled outflow is widely used in hydrologic modeling and has referred to the storage concept (Ponce, 1989). Examples of a reservoir with uncontrolled outflow are a detention pond and an ungauged spillway. In these cases, outflow from a reservoir solely depends on a function of reservoir storage. For hydrologic catchment routing, the outflow is usually assumed to be linearly proportional to the instantaneous storage content. This linear relationship between outflow and storage can be expresses as follows:

$$Q_{out}(t) = \frac{S(t)}{K} \tag{3.5}$$

where

$Q_{out}(t)$	is outflow at time step t [$\mathrm{m}^3\ \mathrm{s}^{-1}$];
$S(t)$	is storage content at time step t [m^3];
K	is storage constant [s].

The change in storage over time interval Δt can be expressed, using Equation (3.6) as

$$\frac{\Delta S(t)}{\Delta t} = K\frac{\Delta Q_{out}(t)}{\Delta t} \tag{3.6}$$

Each linear reservoir has a diffusion effect on an inflow hydrograph. When an inflow hydrograph is routed through a linear reservoir, the diffusion effect decreases a peak of the hydrograph but extends its base time. This diffusion effect is related to the constant K, which is the linear storage coefficient (Equation 3.6). Increasing the value of K results in increased outflow hydrograph diffusion.

Four main discharge components of a hydrologic system – namely surface flow, near-surface runoff (interflow), groundwater flow and delayed groundwater flow (baseflow) can be represented by four linear reservoirs in hydrologic catchment routing. Their storage constants are parameterized based how fast their responses are. The storage constants for surface flow and interflow can be estimated by using an empirical relationship related to watershed characteristics can be derived from average surface slope, flow length and drainage density. The other storage constants for groundwater flow and baseflow are usually determined as constant factors (Riedel, 2004). Afterwards, outcomes of the four linear reservoirs are combined to form runoff hydrograph as depicted in Figure 3.2.

3.2.1.3 Runoff routing

Runoff routing in this section is referred to hydrologic stream channel routing. This hydrologic stream channel routing is a procedure to determine the discharge-time relationship (i.e., the hydrograph) at a point on a drainage network from which the hydrograph was derived or assumed. When a hydrograph travels through a given reach of the channel from upstream to downstream, it is modified in two main processes. Firstly, the effect of channel storage called attenuation redistributes a hydrograph. As a result, the hydrograph peak is diminished and also its shape widens at downstream points. Secondly, the time of the hydrograph peak is lengthened as it passes from upstream to downstream. This process is known as translation (Figure 3.3). In other words, attenuation transforms the shape of the hydrograph, whereas translation changes its position (Chow et al., 1988).

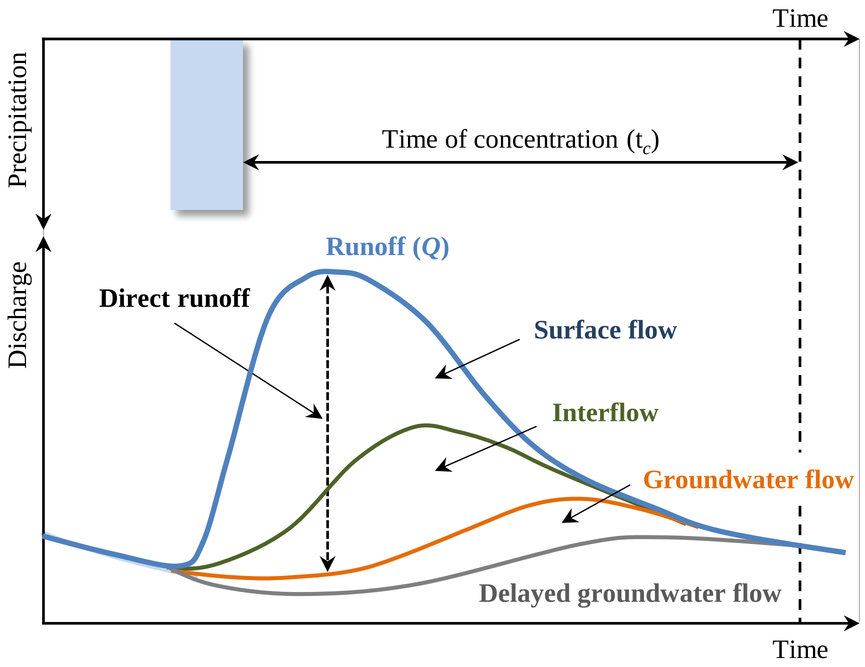


Figure 3.2: Total runoff hydrograph sub-divided into four sub-hydrographs: surface flow, interflow, groundwater and baseflow (modified from Stödter (1994), with additions).

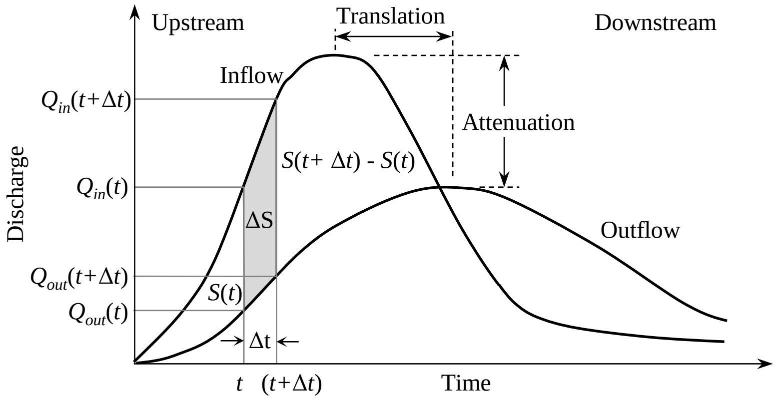


Figure 3.3: Inflow and outflow hydrographs for hydrologic storage routing, and change of storage during a routing period Δt (modified from Chow et al. (1988), with additions).

Hydrologic runoff routing methods based on the storage concept employ the principle of continuity and concentrates on the difference between inflow $Q_{in}(t)$ and $Q_{out}(t)$ outflow rates, which equals to the change in storage S over the time t, i.e.

$$Q_{in}(t) - Q_{out}(t) = \frac{\mathrm{d}S}{\mathrm{d}t} \tag{3.7}$$

Figure 3.3 shows a conceptual interpretation of inflow and outflow hydrographs. The time horizon is divided into intervals of duration Δt. Alternatively, Equation (3.7) can be rewritten in the discrete continuity equation as:

$$\frac{\Delta S}{\Delta t} = \frac{S(t+\Delta t) - S(t)}{\Delta t} = \left(\frac{Q_{in}(t) + Q_{in}(t+\Delta t)}{2}\right) - \left(\frac{Q_{out}(t) + Q_{out}(t+\Delta t)}{2}\right) \tag{3.8}$$

where

$S(t)$	is storage at time step t [m^3];
$S(t+\Delta t)$	is storage at time step $t+\Delta t$ [m^3];
$Q_{in}(t)$	is inflow at time step t [m^3 s^{-1}];
$Q_{in}(t+\Delta t)$	is inflow at time step $t+\Delta t$ [m^3 s^{-1}].
$Q_{out}(t)$	is outflow at time step t [m^3 s^{-1}];
$Q_{out}(t+\Delta t)$	is outflow at time step $t+\Delta t$, [m^3 s^{-1}].

Equation (3.8) states that a change of storage in time ($\Delta S/\Delta t$) balances any difference between inflow and outflow. By substituting the result for $\Delta S/\Delta t$ from Equation (3.8) into Equation (3.6), it forms the inflow-outflow relationship based on the linear storage techniques as given by the following equation (Maniak, 2010).

$$Q_{out}(t+\Delta t) = Q_{out}(t) + C_s\left[Q_{in}(t) - Q_{out}(t)\right] + 0.5C_s\left[Q_{in}(t+\Delta t) - Q_{in}(t)\right] \tag{3.9}$$

where

t	is time t [h];
Δt	is the time intervals [h];
Q_{in}	is inflow [m^3 s^{-1}];
Q_{out}	is outflow [m^3 s^{-1}];
C_s	is coefficient: $C_s = \Delta t \,/\, (K + 0.5\Delta t)$ [-];
K	is storage constant [h].

Translation may be interpreted as the lateral movement of water towards downstream without changing the shape of its wave. Time of translation can be estimated by using the kinematic wave model. For uniform flow, this kinematic wave model assumes that the bottom slope of the channel is parallel to the friction slope and the friction and gravity forces balance each other. This simplification of the kinematic wave model can be described by empirical formulas such as Manning-Strickler or Chezy equations. The Manning-Strickler equation is:

$$V = k_{St} R_{hy}^{2/3} S_f^{1/2} \tag{3.10}$$

The Chezy equation is:

$$V = C_z R_{hy}^{1/2} S_f^{1/2} \tag{3.11}$$

where

V	is the flow velocity [m^2 s^{-1}];
k_{St}	is the Manning-Strickler friction coefficient [m$^{1/3}$ s^{-1}];
C_z	is the Chezy coefficient [m$^{1/2}$ s^{-1}];
R_{hy}	is the hydraulic radius [m].
S_f	is the friction slope approximated by the bottom slope of the channel [m/m].

Actually, there are numerous hydrologic methods to model a hydrologic system of a catchment. Some of them are described above because they have been widely used and have been implemented in a hydrologic modeling system PANTA RHEI that was later used to model a water balance of a study region. The other methods can be referred to the literature (e.g., Chow et al. (1988) and Maniak (2010)).

3.2.2 Rainfall-runoff relations of small and midsize watersheds

3.2.2.1 Rational method

Since the mid-nineteenth century, the Rational method based on a linear rainfall-runoff relation has been widely used for estimating peak discharge on small areas because it does not require historical records (Linsley, 1949; Chow et al., 1988). This Rational method has two main assumptions of uniform storm intensity covering the entire basin and storm duration being greater than or equal to the concentration time (t_c) of the basin. The idea behind the Rational method is that the discharge rate at the outlet will gradually increase until rainfall from the entire basin has traveled to the outlet. At that time, discharge rate also reaches the maximum or equilibrium flow rate (Q_e), which occurs at time t_c.

With respect to the relative magnitudes of concentration time and effective rainfall intensity (I_e), there are three types of basin response (Ponce, 1989). The first type of basin response called as concentrated catchment occurs when the effective rainfall duration (t_r) exactly equals to t_c as illustrated in Figure 3.4a. In this case, the hydrograph reaches its maximum ($Q_{max} = Q_e$) at time t_c when the entire basin is contributing runoff to the outlet. At this time rainfall stops and the hydrograph then begins to recede back to zero. The second type of basin response called as superconcentrated catchment occurs when the effective rainfall duration is greater than t_c. In this case, the runoff reaches its maximum after an elapsed time equal to t_c. As rainfall continues beyond the concentration time, the whole basin continues to contribute to flow at the outlet, and subsequent flows remain constant at its equilibrium value. After the rainfall has stopped, the runoff gradually recedes back to zero. An example hydrograph for this type of basin response is sketched in Figure 3.4b. The third type of basin response called as subconcentrated catchment occurs when the effective rainfall duration is shorter than t_c. The hydrograph for this case as shown in Figure 3.4c does not reach the equilibrium value. After rainfall stops, the hydrograph gradually recedes back to zero.

3.2.2.2 SCS method

The Soil Conservation Service (SCS) method, developed by the United States Department of Agriculture (USDA) in 1954, is documented in USDA (1986). This method was chosen for estimating direct runoff in this dissertation. This method is widely used for estimating floods on small to medium-sized ungauged basins in the US and many other countries (Pilgrim and Cordery, 1993; Thompson, 1999; Maniak, 2010). The basic theory of the SCS method is based on the water balance equation of four parameters: precipitation, direct runoff, initial abstraction and cumulative infiltration, which can be expressed as:

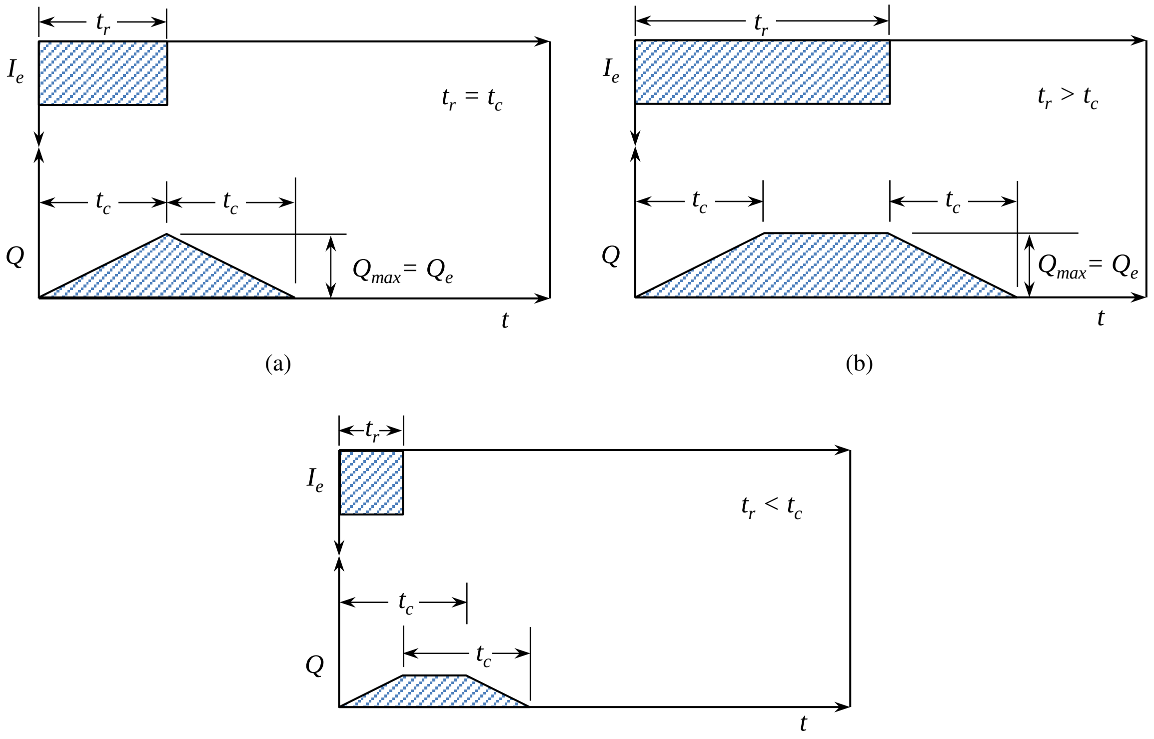
$$F = P - I_a - P_e \tag{3.12}$$

where

F is cumulative infiltration [mm];
P is the depth of precipitation [mm];
I_a is the initial abstraction [mm];
P_e is the depth of effective precipitation or direct runoff [mm].

The hypothesis of the SCS method is that the ratio of actual amount of direct runoff (P_e) to potential runoff ($P - I_a$) and the ratio of actual infiltration (F) to the potential maximum retention (S_e) are equal (Chow et al., 1988). This hypothesis can be rewritten as:

$$\frac{P_e}{P - I_a} = \frac{F}{S_e} \quad \text{when } P > I_a. \tag{3.13}$$



(a)

(b)

(c)

Figure 3.4: Runoff concentration (taken from Ponce (1989), modified).

Substituting F from Equation (3.12) into Equation (3.13) gives the basic equation for calculating the direct runoff from a storm as follows:

$$P_e = \frac{(P - I_a)^2}{P - I_a + S_e} \quad \text{when } P > I_a. \tag{3.14}$$

The parameter S_e representing the potential maximum retention depends upon the soil, vegetation, land use and antecedent precipitation conditions of the basin. For convenience in practical application, the potential retention can be expressed in terms of a dimensionless curve number (*CN*) as:

$$S_e = 254 \left(\frac{100}{CN} - 1 \right) \tag{3.15}$$

where the potential retention S_e is in millimeters.

The curve number depends upon the soil, land use/cover and antecedent precipitation conditions of the drainage basin and is given in forms of tables, which can be found in e.g., Chow et al. (1988), Pilgrim and Cordery (1993) or Maniak (2010). For the purpose of dealing with the time-dependent soil moisture in long-term simulations, the SCS-CN method was extended by modifying *CN* according to soil moisture conditions (Riedel, 2004). The modified *CN* is so called “the relative soil moisture curve number” (CN_{smd}). The CN_{smd} is the relationship between the soil moisture content and the *CN* and is represented by the following equation:

$$CN_{smd} = \frac{1000}{\frac{1000}{CN} - \frac{SM}{25.4}} \qquad \text{for } CN < 100 \tag{3.16}$$

$$CN_{smd} = 100 \qquad \text{for } CN = 100 \tag{3.17}$$

where

CN_{smd}	is the relative soil moisture curve number [-];
SM	is the relative soil moisture [mm].

The relative soil moisture parameter (SM) depends on precipitation, actual evaporation and current soil moisture (Stödter, 1994). It indicates the difference between the field capacity and current soil moisture (Riedel, 2004). The initial value of the relative soil moisture is given as a calibration parameter. During rainy periods, the soil moisture increases because of water infiltration. However, the decrease of the relative soil moisture is as a consequence of runoff formation and the actual evaporation. The relative soil moisture parameter at time *t* can be expressed as:

$$SM(t) = SM(t-1) + P_e - q_p(t) + q_{SM}(t) - E(t) \tag{3.18}$$

where

$SM(t)$	is the relative soil moisture [mm];
$SM(t-1)$	is the relative soil moisture from previous time step [mm];
P_e	is the depth of effective precipitation or direct runoff [mm];
$q_p(t)$	is the depth of areal runoff [mm].
$q_{SM}(t)$	is the saturation of soil moisture from formerly effective precipitation [mm].
$E(t)$	is the actual evaporation [mm].

3.2.3 Issues of space and time scales in hydrology

Hydrologic processes occur at a broad range of scales such as from unsaturated flow in a few meters of soil to floods in river systems covering a million square kilometers. In addition, hydrologic processes occur within a few minutes duration like flash floods and over a hundred years like flow in aquifers (Blöschl and Sivapalan, 1995). In hydrology and water resources management, scientists often investigate and analyze surface hydrologic processes at the scale of watersheds. Watersheds greater than 10^5 km^2 are normally considered as large watersheds and are categorized as macroscale (Figure 3.5). Watersheds varying from 100 to 5000 km^2 are defined as midsize watersheds (Ponce, 1989). Watersheds smaller than 100 km^2 are considered as small watersheds.

Land surface hydrology between large and small watersheds deviates in some aspects due to the variety of landscape forms such as the diversity of soil types, topography, land uses, and vegetation. Moreover, there are differences with regards to the availability of data between large and small watersheds. For instance, observation of water levels for larger rivers is usually available for longer periods, whereas water-level observation for small rivers may be available on the event basis or may be unavailable at all. Thus, approaches applied to larger watersheds may be different from ones applied to small watersheds (Liebscher, 1993).

Figure 3.5 presents the areas covered by hydrologic models in space and time scales. From this figure, it can be seen that flood forecasting models and water balance models are preferably applicable to large watersheds if long time series of data exist. In addition, with large time steps such as 30 years water balance models can be applied to nearly all sizes of watersheds. In contrast, at small time steps such as hourly and daily may require to model a hydrologic system in small watersheds.

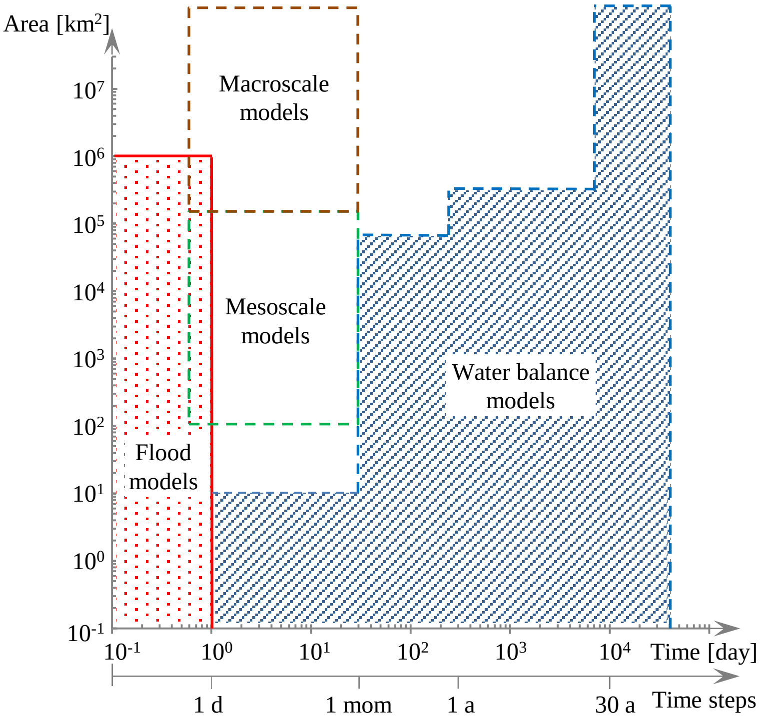


Figure 3.5: Space and time scales in hydrologic modeling (after Liebscher (1993), modified).

3.3 Hydrodynamic modeling

The flow processes in natural watercourses are complex. However, in general, several commonly terms of the flow can be treated by numerical methods. This section describes some of the numerical methods used in hydrodynamic models. The hydrodynamic models are used to convert a volume of water and then to simulate water levels and overland flows. Basis information required for the models is spatially detailed topography, flow resistances, hydraulic boundary conditions and initial flow conditions.

Hydrodynamic models fundamentally involve the computation of flood wave propagation. Under certain simplifying assumptions, this computation can be governed by either a one-dimensional equations or the frame of a two-dimensional analysis.

3.3.1 One-dimensional hydrodynamic models

The one-dimensional (1D) computation of flow propagation in a natural watercourse reduces multidimensional and complex flow processes to a 1D level by adopting a constant velocity and movement of the water along the watercourse. Most of 1D hydrodynamic models approximate the flow based on the numerical solution of the Saint-Venant equations. The basic assumptions of these equations are given as follows:

- The flow is one-dimensional. It means that the cross-sectional velocity is uniform and the water level across the cross-section is horizontal.
- The direction of the flow is vertical to the selected cross-sections.

- The streamline curvature is small.
- The pressure distribution is hydrostatic so that vertical accelerations are negligible.
- The effects of boundary friction and turbulence can be accounted for through resistance laws comparable to those used for steady-state flow.
- The average channel bed slope is small so that the cosine of the angle can be replaced by unity.

The Saint-Venant equations (1871) for 1D unsteady flow are two coupled partial differential equations, which describe mass and momentum between adjacent cross-sections. The first equation representing the mass conservation can be expressed by Equation (3.19) while the second equation corresponds to the conservation of momentum, which is expressed by Equation (3.20).

$$\frac{\partial Q}{\partial x} + \frac{\partial A_x}{\partial t} = 0 \tag{3.19}$$

$$\underbrace{\frac{\partial Q}{\partial t}}_{I} + \underbrace{\frac{\partial}{\partial x}\left(\alpha \frac{Q^2}{A_x}\right)}_{II} + \underbrace{\frac{gQ \cdot |Q|}{C_z^2 A_x R_{hy}}}_{III} + \underbrace{gA_x \frac{\partial h}{\partial x}}_{IV} = 0 \tag{3.20}$$

where

Q	is the discharge [$\mathrm{m^3\,s^{-1}}$];
A_x	is the cross-sectional area of the flow [$\mathrm{m^2}$];
x	is the space coordinate in flow direction [m];
t	is time [s];
α	is the Boussinesq coefficient [-].
g	is the acceleration due to gravity [$\mathrm{m\,s^{-2}}$];
C_z	is the Chezy coefficient [$\mathrm{m^{1/2}\,s^{-1}}$].
R_{hy}	is the hydraulic radius [m].
h	is the water depth [m].

Equation (3.20) includes acceleration, hydrostatic gradient, bed friction and gravity terms, which are capable for simulating dynamic waves. The combination of the part *III* and *IV* in this equation is considered as a diffusive wave model. It consists of hydrostatic gradient, bed friction and gravity terms. The part *IV* of Equation (3.20) is considered as a kinematic wave model (Chow, 1959; Chow et al., 1988; Henderson, 1966; Montero et al., 2013).

3.3.2 Two-dimensional hydrodynamic models

Two-dimensional (2D) hydrodynamic models are generally based on the depth-averaged Navier-Stokes equations, also called shallow-water equations. The shallow-water equations describe a thin thickness of water over terrain. General assumptions of these shallow-water equations are (i) incompressible and homogenous fluid; (ii) pressure distribution is hydrostatic; (iii) friction is ignored and velocity distributions in the vertical are assumed to be uniform; and (iv) average bottom slope is small. These assumptions simplify the global equations to only three equations: the continuity equation and the momentum conservation equations in direction of the *x*- and *y*-axis (Steffler and Blackburn, 2002; Pasche et al., 2006). Figure 3.6 illustrates shallow-water flow.

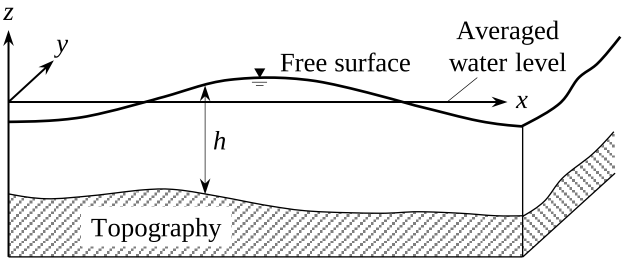


Figure 3.6: Definition sketch for the shallow-water equations.

The continuity equation is given by Equation (3.21).

$$\frac{\partial h}{\partial t} + \frac{\partial\left(\overline{u}h\right)}{\partial x} + \frac{\partial\left(\overline{v}h\right)}{\partial y} = 0 \tag{3.21}$$

where

h	is water depth [m];
t	is time [s];
$\overline{u}$	is depth-averaged velocity in direction of the *x*-axis [$\mathrm{m\,s^{-1}}$];
$\overline{v}$	is depth-averaged velocity in direction of the *y*-axis [$\mathrm{m\,s^{-1}}$];
x and *y*	are distance in the horizontal direction [m].

The momentum equations in direction of *x*- and *y*-axis are given as follows:

$$\begin{aligned}\frac{\partial\overline{u}}{\partial t} + \overline{u}\frac{\partial\overline{u}}{\partial x} + \overline{v}\frac{\partial\overline{u}}{\partial y} = {} & -g\frac{\partial\left(z_0+h\right)}{\partial x} + \frac{1}{h}\frac{\partial}{\partial x}\left[h\left(\frac{1}{\rho_w}\tau_{xx} - \overline{\widetilde{u}\,\widetilde{u}}\right)\right] + \frac{1}{h}\frac{\partial}{\partial y}\left[h\left(\frac{1}{\rho_w}\tau_{xy} - \overline{\widetilde{u}\,\widetilde{v}}\right)\right] \\ & + \frac{1}{h\rho}\left(\tau^{x}_{wind} - \tau_{so,x}\right) + 2\omega v\sin\theta\end{aligned} \tag{3.22}$$

$$\begin{aligned}\frac{\partial\overline{v}}{\partial t} + \overline{u}\frac{\partial\overline{v}}{\partial x} + \overline{v}\frac{\partial\overline{v}}{\partial y} = {} & -g\frac{\partial\left(z_0+h\right)}{\partial y} + \frac{1}{h}\frac{\partial}{\partial x}\left[h\left(\frac{1}{\rho_w}\tau_{yx} - \overline{\widetilde{u}\,\widetilde{v}}\right)\right] + \frac{1}{h}\frac{\partial}{\partial y}\left[h\left(\frac{1}{\rho_w}\tau_{yy} - \overline{\widetilde{v}\,\widetilde{v}}\right)\right] \\ & + \frac{1}{h\rho_w}\left(\tau^{y}_{wind} - \tau_{so,y}\right) + 2\omega u\cos\theta\end{aligned} \tag{3.23}$$

where

g	is the acceleration due to gravity [$\mathrm{m\ s^{-2}}$];
z_0	is ground surface elevation [m];
ρ_w	is the water density [$\mathrm{kg\ m^{-3}}$];
τ_{xx}	is the deviatoric stress acting in *x*-direction on the *x*-face of control volume [$\mathrm{N\ m^{-2}}$];
τ_{xy}	is the deviatoric stress acting in *x*-direction on the *y*-face of control volume [$\mathrm{N\ m^{-2}}$];
τ_{yx}	is the deviatoric stress acting in *y*-direction on the *x*-face of control volume [$\mathrm{N\ m^{-2}}$];
τ_{yy}	is the deviatoric stress acting in *y*-direction on the *y*-face of control volume [$\mathrm{N\ m^{-2}}$];
τ_{wind}	is the wind shear stress with superscripts *x* and *y* indicate *x*- and *y*-direction, respectively, [$\mathrm{N\ m^{-2}}$];
τ_{so}	is the bottom shear stress with superscripts *x* and *y* indicate *x*- and *y*-direction, respectively, [$\mathrm{N\ m^{-2}}$];

$\widetilde{u}$	is a deviation of the depth-averaged velocity in x direction [m s^{-1}];
$\widetilde{v}$	is a deviation of the depth-averaged velocity in y direction [m s^{-1}];
ω	is the angular rotation rate [radians hour^{-1}];
θ	is latitude [°].

Equations (3.22) and (3.23) may be rewritten in a tensor form as

$$\underbrace{\frac{\partial \overline{u_i}}{\partial t}}_{1} + \underbrace{\overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}}_{2} = \underbrace{-g\frac{\partial\left(z_0 + h\right)}{\partial x_i}}_{3} + \frac{1}{h}\frac{\partial}{\partial x_j}\left[h\left(\frac{1}{\rho}\underbrace{\tau_{ij}}_{4} - \underbrace{\overline{\widetilde{u_i}\,\widetilde{u_j}}}_{5}\right)\right] + \underbrace{\frac{1}{h\rho}\left(\tau_{wind,i} - \tau_{so,i}\right)}_{6} + a_{F_i} \tag{3.24}$$

where

a_{F_i}	is acceleration component due to the Coriolis force,[m s^{-2}];
i, j	is 1, 2 stands for the *x*- and *y*-direction, and runs the Einstein sum convention [-].

The first term in Equation (3.24) represents the rate of volume change over time. The second term refers to the convective momentum transport. The third term accounts for the gravitational force. The fourth term is the diffuse momentum transport. The fifth term describes the dispersive momentum transport. For uniform flows this term can be omitted. The sixth term incorporates the outside forces, which are wind shear stress and bottom shear stress. In most of cases, the wind shear stress can be omitted. The bottom shear stress can be estimated using an empirical equation as follows:

$$\tau_{so} = c_f \rho\, u_i \sqrt{u_i^2 + u_j^2} \tag{3.25}$$

where c_f is the friction coefficient, which can be determined using either the flow laws of Darcy-Weisbach or Gauckler-Manning-Strickler as given in Equations (3.26) and (3.27), respectively.

$$c_f = \frac{\lambda_D}{8} \tag{3.26}$$

$$c_f = \frac{\rho g}{k_{St}^2 h^{1/3}} \tag{3.27}$$

where

λ_D	is Darcy-Weisbach coefficient [-];
k_{St}	is Manning-Strickler coefficient [m$^{1/3}$ s^{-1}].
ρ	is the water density [kg m^{-3}];
g	is gravitational acceleration [m s^{-2}];
h	is water depth [m].

In short, Equations (3.21) and (3.24) are the conservative form of the shallow water equations. It is commonly assumed that the diffusion, the dispersion, wind shear stress and Coriolis force terms are negligible.

3.4 Statistical analyzes

3.4.1 Principal component analysis (PCA)

PCA is a widely used multivariate technique to discover patterns in large data sets. Here, PCA is selected because of its robustness. PCA can form relationships between variables and automatically weight these relationships by their variance, without involving "expert" judgment.

The particular purpose of PCA is to interpret linear relationships between variables in both differences and similarities (Jolliffe, 2002). For this purpose, the maximum variance from the multivariate data sets is extracted and new sets of variables are then formed. The new sets of variables called the principal components (PCs) are linear combinations of the initial data sets. Each subsequent principal component captures the greatest possible variance in the initial variables and is strongly correlated within one group of variables. Nonetheless, it is always uncorrelated with one another principal component. All PCs are arranged in order of decreasing variance. The first principal component accounts for the greatest percentage of the total variance, whereas the last accounts for the least percentage (Dunteman, 1989).

Suppose that there are just two variables: X_1 and X_2, for example. Their 1000 sample data randomly generated by using MATLAB software are plotted in X_1 and X_2 axes as shown in Figure 3.7a. As mentioned above, the first principal component (PC1) captures the maximum variability of the data. The second principal component (PC2) is orthogonal to the first one. Each principal component is a linear transformation of the original data. These original data can be rotated by PCA technique to new coordinates on PC1 and PC2 plane as shown in Figure 3.7b.

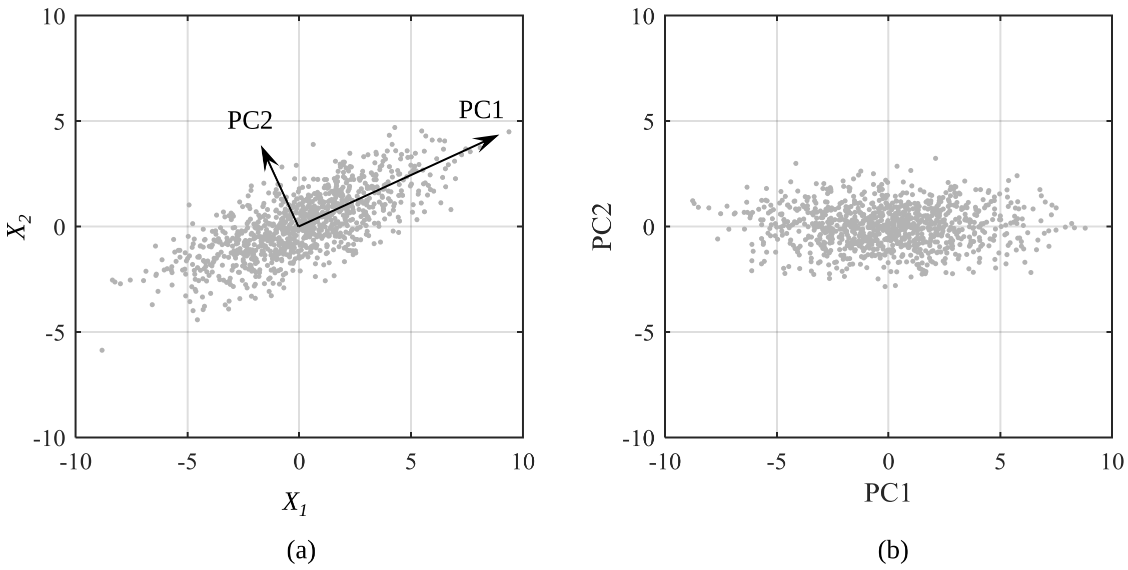


Figure 3.7: Scatter plot of data randomly created by MATLAB: (a) original data on X_1 and X_2 plane and (b) transformed data on PC1 and PC2 plane.

If the number of variables is greater than two, it is useful to perform PCA with assistance of matrix and linear algebra operations. Suppose that *X* is an *n x m* data matrix whose rows correspond to the number of observations and columns correspond to the number of variables. The new sets of variables are formed as linear combinations of the initial ones. Linear combinations (principal components) of the X can be defined as follows:

$$\begin{aligned} PCi &= a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{im}X_m \\ &= a_i^T X \end{aligned} \tag{3.28}$$

where

PCi	is a vector of scores for the i^{th} principal component with $i = 1, 2, \cdots, m$;
m	is the number of variables;
a_i^T	consists of $a_{i1}, a_{i2}, \cdots, a_{im}$, which are the coefficients of the i^{th} principal component.

In addition, these coefficients are weights for each variable so that the sums of their squared coefficients add up to one. The coefficients of all PCs can be expressed in matrix *U*. In other words, *U* is a matrix combining the eigenvectors of matrix C_x, the variance-covariance matrix of *X*. By definition, this matrix C_x is an *m* x *m* diagonal matrix which describes the covariance between variables. This measures how much each variable varies from the mean with respect to each other variable and how two variables increase together.

Afterwards, the eigenvectors and eigenvalues of C_x can be algebraically computed using the matrix equation:

$$C_x = U \Lambda U^T \tag{3.29}$$

where

C_x	is a covariance matrix of the *X*, *Cov(X)*;
U	is a matrix combining eigenvectors of C_x;
Λ	is a matrix combining eigenvalues of C_x;

Maritx Λ is a diagonal matrix, which can be expressed as follows:

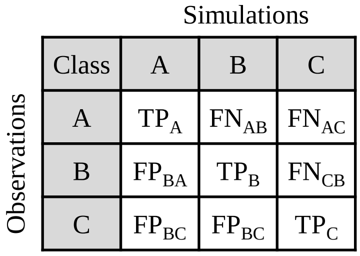
$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \end{bmatrix} \tag{3.30}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_m$ are eigenvalues of $PC1, PC2, \cdots, PCm$, respectively.

PCA searches the eigenvalues of the variance-covariance matrix of *X* (i.e., *Cov(X)*). The first eigenvector is associated with the largest eigenvalue, λ_1. To put it another way, the first principal component (PC1) captures the majority of the variance in the data sets. The second principal component (PC2) which is a linear combination perpendicular to and unrelated to the PC1 has the second-largest eigenvalue, λ_2 and so forth.

3.4.2 Evaluation of classification performance

A confusion matrix, also called an error matrix (Congalton, 1991), is an effective way of summarizing and visualizing the performance of a classification model (or "classifier"). For example, Figure 3.8 shows a table demonstrating comparison results among different classes of observations and model simulations. The number of rows and columns represents the number of classes in the observations and model simulations, respectively. The diagonal of the table indicates true positives (TP), the number of both the observations and simulations matched in the same class. Another case when a model result indicates that a given condition exists in a certain class, but it does not, is called as false positives (FP). On the contrary, false negatives (FN) are when a model result indicates that a given condition does not exist but it does.



Simulations

	Class	A	B	C
Observations	A	TP_A	FN_{AB}	FN_{AC}
	B	FP_{BA}	TP_B	FN_{CB}
	C	FP_{BC}	FP_{BC}	TP_C

Note:
- True positives (TP)
- False positives (FP)
- False negatives (FN)

Figure 3.8: Confusion matrix showing comparisons between observations and simulations.

Agreement ratio

The overall accuracy of a classification performance can be estimated by comparing the classified values of observations with those of simulations. An accuracy, hereafter called "agreement ratio", is the proportion of correctly classified observations over all observations (Congalton, 1991). This results in the range from zero to one where values close to one indicates a high degree of agreement. An agreement ratio of the confusion matrix in Figure 3.8 can be computed as follows:

$$Agreement\ ratio = \frac{TP}{TP + FP + FN} \tag{3.31}$$

where

TP	is true positives;
FP	is false positives;
FN	is false negatives.

F-score

F-score usually defined as the harmonic mean of precision and recall is a measure of a classification performance. Precision indicates how many relevant results are correctly classified by a classifier, whereas recall represents how many relevant results are retrieved (Van Rijsbergen, 1975). The F-score reaches its best at one and worst score at zero. The formula of the F-score can be expressed as follows:

$$F - score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{3.32}$$

where

$precision$	is the number of true positives divided by the number of true positives and false positives, also written as $TP/(TP + FP)$;
$recall$	is the number of true positives divided by the number of true positives and false negatives, also written as $TP/(TP + FN)$;

3.5 Flash flood hazard mapping

Generally, the hazard of flooding is usually defined in terms of its probability of occurrence and its magnitude (Kron, 2002). To estimate and map flood hazard, several approaches can be applied according to available data. In this section, four commonly used approaches to produce flood hazard maps (and flash flood hazard maps) are described and shown in Figure 3.9.

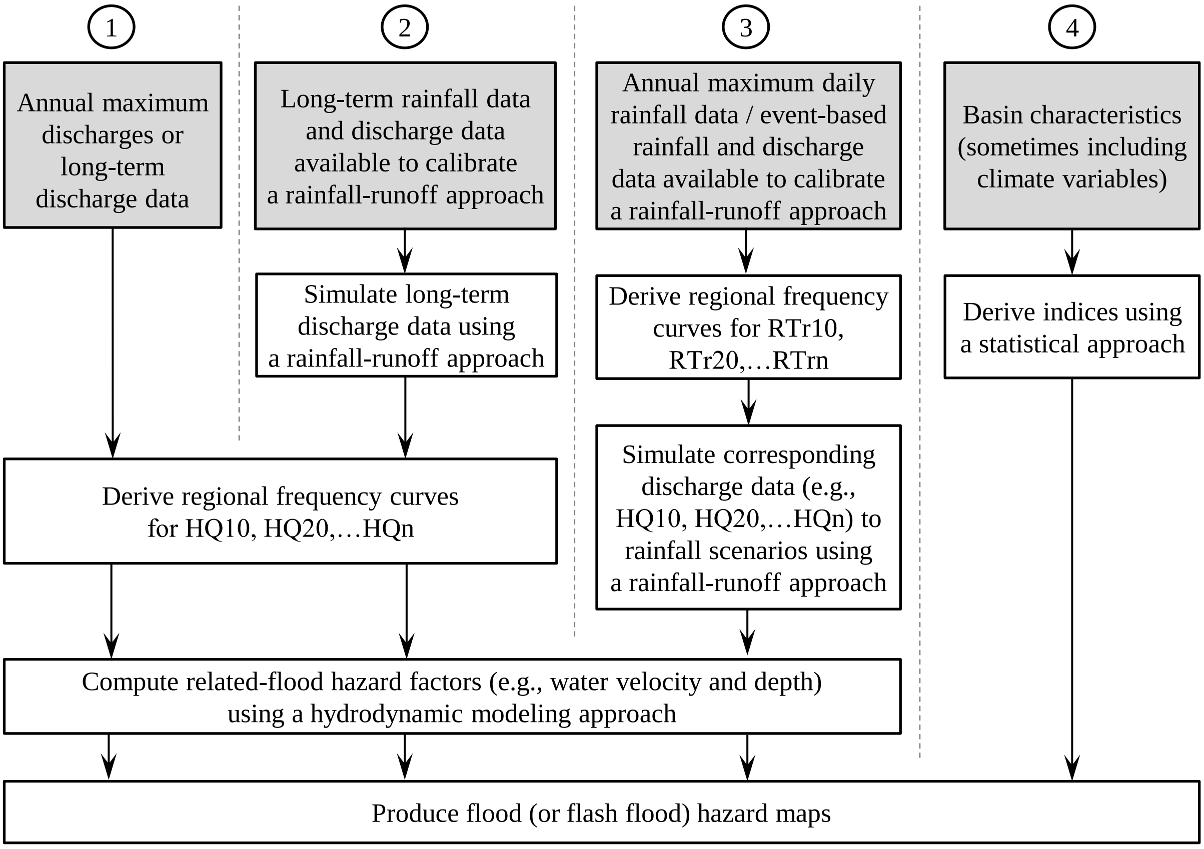


Figure 3.9: Four different approaches to produce flood hazard maps, depending on available data (note: HQn is an n-year return period flood and RTrn is an n-year return period rainfall event).

The first approach is suitable for areas where annual maximum discharges or long-term (i.e., about 30 years) discharge data are available and consists of three steps. In the first step, these data are often employed to derive regional frequency curves that represent the probability of an event's occurrence of given magnitude (Chow et al., 1988). The regional frequency curves are usually used to estimate the design flood, for example, a 10-year return period flood (HQ10) (Di Baldassarre, 2012). In the second step, the design flood is used as input to a hydrodynamic modeling approach in order to simulate flood inundation extent and related-flood hazard factors (e.g., floodwater velocity and depth). In the third step, related-flood hazard factors – namely, floodwater velocity and depth, are usually used to assess the flood hazard degree for people through an empirical formula (e.g., Ramsbottom et al., 2006) and to produce flash flood hazard maps.

For the second approach, computational steps to estimate the flood hazard degree are almost comparable to the first approach. However, the second approach simulates long-term discharge data based on long-term rainfall data by using a rainfall-runoff approach. Once the long-term discharge data are obtained, it can be followed the computational three steps of the first approach to produce flash flood hazard maps.

The third approach can be applied to areas where long-term time series of (sub-daily or daily) rainfall data are not available, but annual maximum daily rainfall data are available to derive regional frequency curves of rainfall. These regional frequency curves of rainfall can be used as the design rainfall, for example, a 10-year return period extreme rainfall event (RTr10). Moreover, this third approach requires to have event-based rainfall and discharge data in order to construct rainfall-runoff relationships. These rainfall-runoff relationships that are usually formed by using a rainfall-runoff approach can be used to simulate the design flood (e.g., HQ10) from the design rainfall (e.g., RTr10). After the design flood is obtained, the following step is similar to the third step of the first approach.

The last approach (i.e., index-based approach) is in contrast to the above-mentioned approaches. It derives indices (e.g., flash flood potential indices, flash flood susceptibility) from basin characteristics and sometimes includes climate variables by using a statistical approach (e.g., Minea, 2013; Khosravi et al., 2016; Costache and Zaharia, 2017). This statistical approach is mainly used to analyze variables related to basin characteristics and weight these variables in order to construct the indices. To determine the degree of flash flood hazard, these indices need to be categorized into several classes (e.g., very low, low, moderate, significant, and extreme). Afterwards, their corresponding areas are defined flash flood hazard zones.

In brief, the first three approaches can be applied to produce flash flood hazard maps corresponding to a certain return period based on the design flood. Hence, different flash flood hazard maps under different exceedance probabilities can be obtained. On the contrary, a flash flood hazard map produced by the last approach (i.e., index-based approach) can be usually used to represent only a single flood event.

In this dissertation, the latter two approaches were used to assess flash flood potential in the study regions (as described later in Chapter 4) because of poor data availability. The application of these two approaches is presented later in Chapter 6. However, the index-based approach in this research work is new and has the uniqueness that it can produce flash flood hazard maps corresponding to different rainfall scenarios (see Chapter 7).

[This page intentionally left blank]

4 Study Area and Data Collection

On the basis of an extensive pre-investigation of possible case study areas, two study areas (the upper parts of the Nan River and the Ping River basins as shown in Figure 4.1) were finally chosen because of flash flood occurrence in the past and suitability for transfer of results. Past flash flood events in many provinces of Thailand have been recorded in reports[1] on landslide inventories (in local language) contributed by the Department of Mineral Resources (DMR) since 2012. Although the reports have focused on landslide inventories and data gathering, they also provide useful information about flash floods such as historic flood events and sub-districts affected by these events in the past. In the reports, devastating flash floods frequently occur in many areas of the selected two study areas. Furthermore, the upper Nan River and the upper Ping River basins lie in the northern Thailand and share a similar climate. Therefore, results and flash flood concepts obtained from one were expected to be transferable to the other.

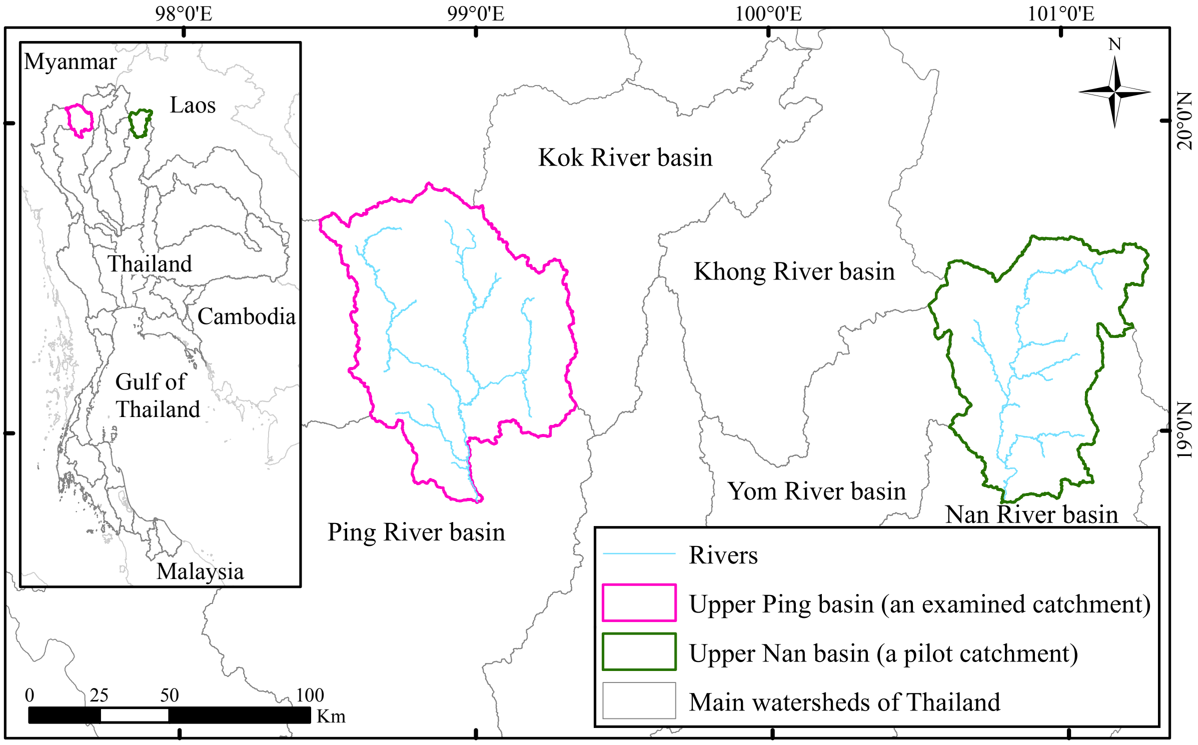


Figure 4.1: Study area: the upper part of the Nan River basin (a pilot catchment) and the upper part of the Ping River basin (an examined catchment).

As mentioned above, two study areas were the upper Nan River and the upper Ping River basins. Their rivers are two principal tributaries of the Chao Phraya River and flow from north to south. At the confluence

---

[1] Available on the website: http://www.dmr.go.th/main.php?filename=index (accessed on December 23, 2017)

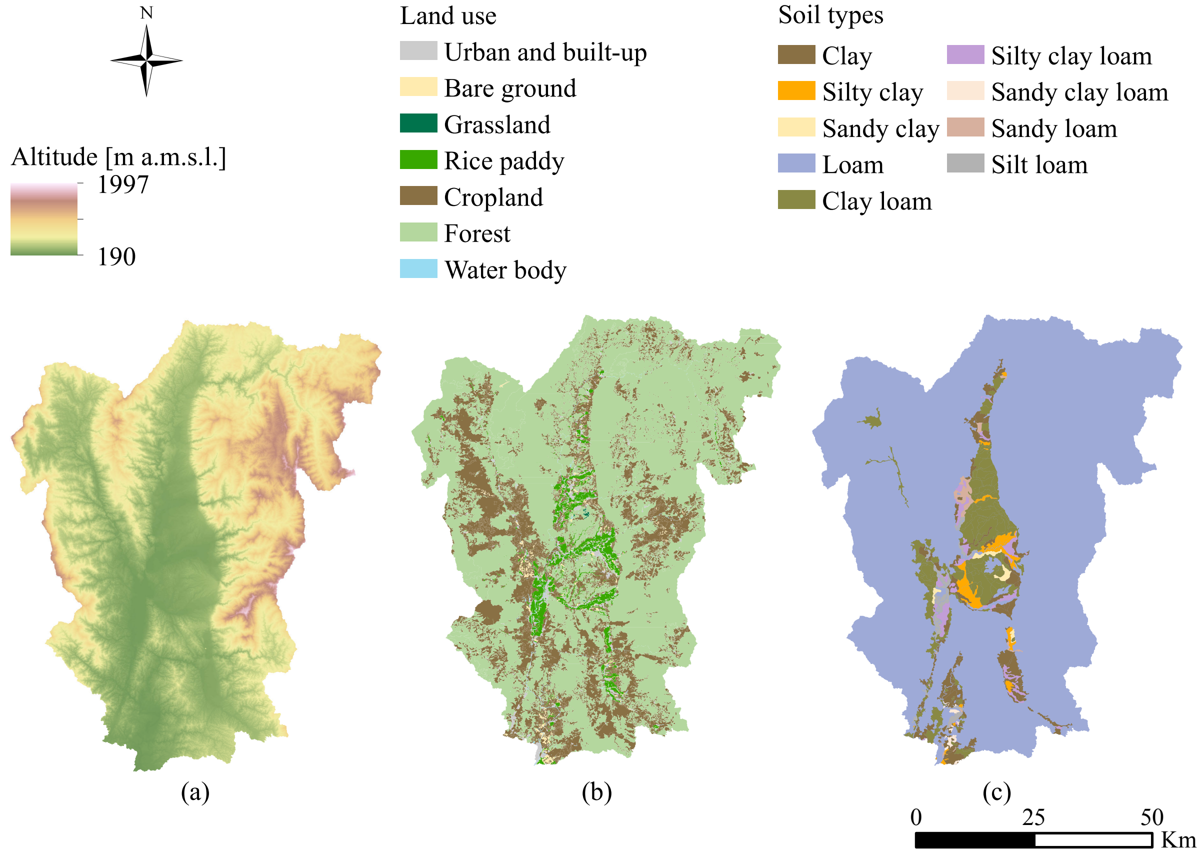
of these rivers in Nakhon Sawan province the Chao Phraya River forms. The Chao Phraya River drains through the central plain, where ancient capitals like Bangkok and Phra Nakhon Si Ayutthaya are located, and exits to the Gulf of Thailand.

The Nan River and Ping River basins have a tropical climate influenced by the South Asian Monsoon. The climate is prevailed by the annual monsoon cycle, which alternates dry and wet seasons. The weather is almost dry between November and early April. In late April, the wet seasonal monsoon arrives in those regions and usually withdraws from them by October. This tropical monsoon carries moist air from the Indian Ocean and brings ample rains to the regions. As a result, both the Nan River and Ping River basins are affected by “normal” floods, but also by flash floods almost every year because of intense rains in short periods or prolonged moderate rainfall and their own catchment characteristics. Generally, downstream areas have an almost flat landscape while the headwater basins have steep slopes. When heavy rainfall induces flash flooding in upstream basins, the lower basin might have high potential for a long-period flooding. As flash floods are the main focus here, the upper Nan River basin was chosen as a pilot catchment. The upper Ping River basin was carried to examine the findings obtained from the upper Nan River basin.

4.1 Upper Nan River basin: a pilot region

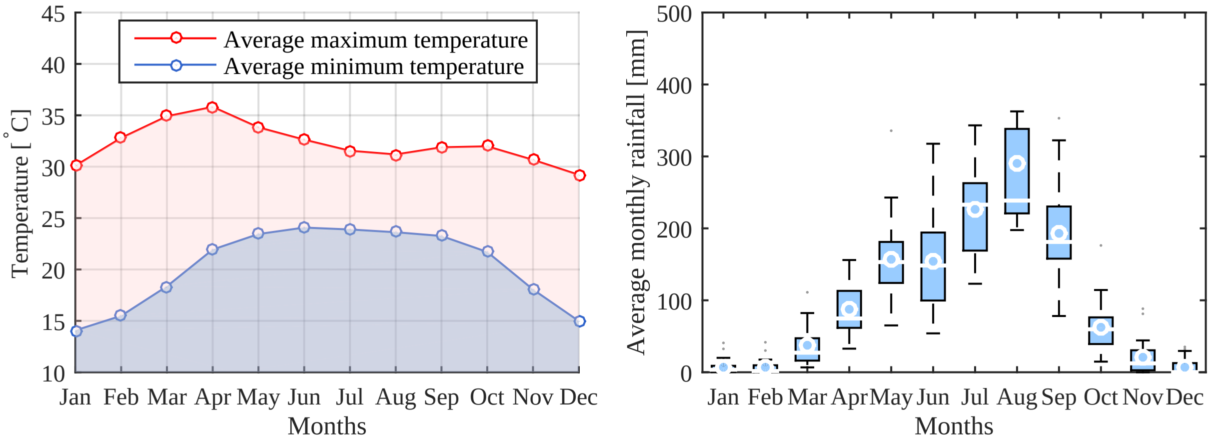
The upper Nan River basin used as a pilot catchment covers an area of 4609 km^2, which is about 13 per cent of the total Nan River basin area. This pilot catchment is surrounded by two main mountainous ranges, which are the Luang Prabang Range in the northern and eastern parts and the Phlueng Range in the west. Its relief varies from 1997 m to 190 m a.m.s.l. (Figure 4.2a). The Nan River and its tributaries rise from those ranges. The main river flows southwards. Its plain includes the city of Nan Province, an important tourism center of Thailand. There are many historic buildings settled along the Nan River as well as densely local communities. The urban and built-up areas cover about 1% of the study area. However, the predominant land use for this catchment is forest, which covers approximately 78% of the total area based on survey information in 2009 obtained from the Land Development Department (LDD) in Thailand (Figure 4.2b). According to the LDD, the forest areas in the catchment have been decreased because of deforestation. The catchment is mainly formed by loam and clay loam (Figure 4.2c). These dynamic land use situation, especially the deforestation, and soil properties may increase the potential to flash flooding.

Flash flood occurrence within the pilot catchment is the consequence of three driving factors: meteorology, geomorphology and hydrology. The climate of this catchment is controlled by tropical monsoons. Figure 4.3a shows the average monthly temperature of the catchment based on the records in 1993–2012 from four meteorological stations: Muang Nan, Nan Agrometeorological, Tha Wang Pha and Thung Chang stations. The average annual temperature is about 26°C. The average maximum and minimum monthly temperatures are 32°C and 20°C, respectively. As a result of this regional climate setting, the catchment receives the massive majority of its annual rainfall during the wet season between late April and early October, which is responsible for approximately 90% of total annual rainfall (Figure 4.3b). On average, there are one or two tropical depressions per year across the catchment. These bring heavy rains in a few hours and continuously moderate rains (up to 15 $\text{mm}\,\text{h}^{-1}$) for a few days, which usually induce flash floods in small sub-basins located in the headwater and mountainous areas. The steep surfaces of these sub-basins lead to quick hydrologic response when they receive such tropical rainstorms.



(a) (b) (c)

Figure 4.2: The upper Nan River basin: (a) topography of upper Nan River basin (source: RID), (b) land use map based on survey information in 2009 (source: LDD), and (c) soil map (source: LDD).



(a) (b)

Figure 4.3: Average monthly temperature and rainfall observations in the period 1993–2012 for the upper Nan River basin: (a) plot for average maximum temperature (red solid line) and average minimum temperature (blue solid line) and (b) box and whisker plot for average monthly rainfall.

4.1.1 Meteorological data

4.1.1.1 Ground based meteorological observations

Ground based meteorological observations have been primarily used in this dissertation. Available meteorological variables with daily time-series were obtained from the Thai Meteorological Department (TMD) as listed in Table A.1 (in Appendix A.1.1). The variables are rainfall, temperature, sunshine duration, relative humidity, evaporation, and wind speed. Rainfall data recorded every 24 h with standard raingauges is available from 1993 to 2012. One of 11 raingauges is operated by the Department of Water Resources (DWR). Another type of raingauge automatically measures rainfall every 15 min. Since 2006, automatic raingauges have been in operation and have belonged to DWR. A list of these raingauges used in this dissertation can be found in Table A.2 (in Appendix A.1.1). Many of them had been operated for only three or four years and contained too little heavy rainfall data for statistical analysis. Locations of raingauges are indicated by the plus (+) symbol for daily records and the circle (o) symbol for15-min records as shown in Figure 4.4.

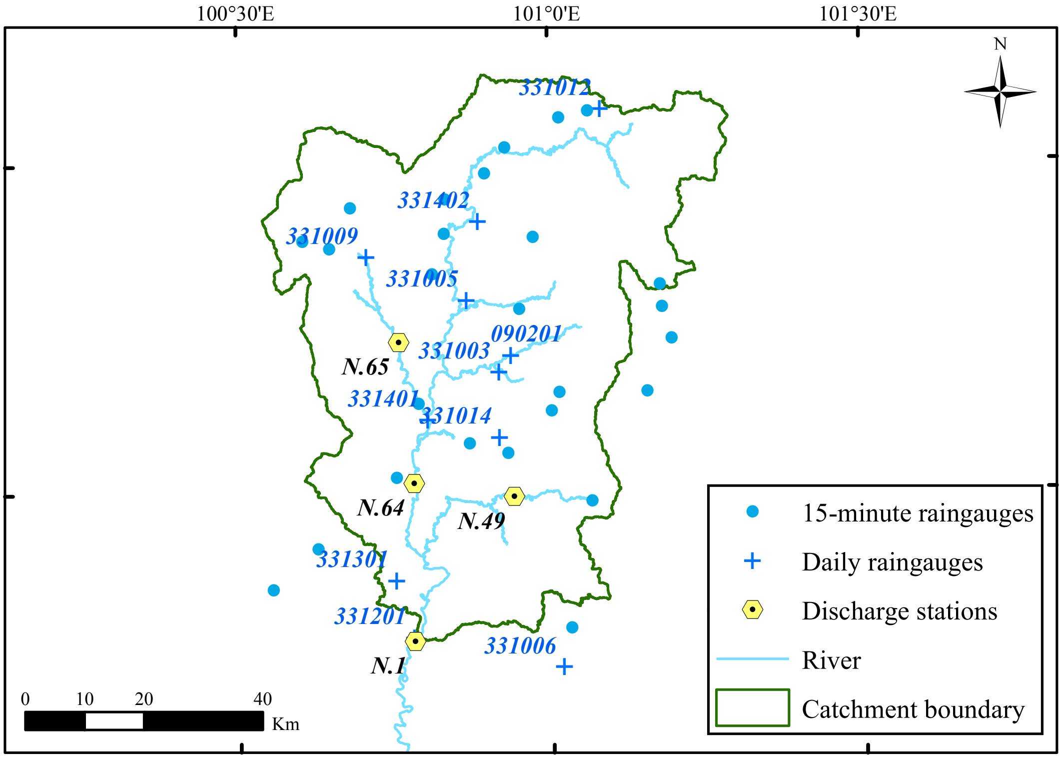
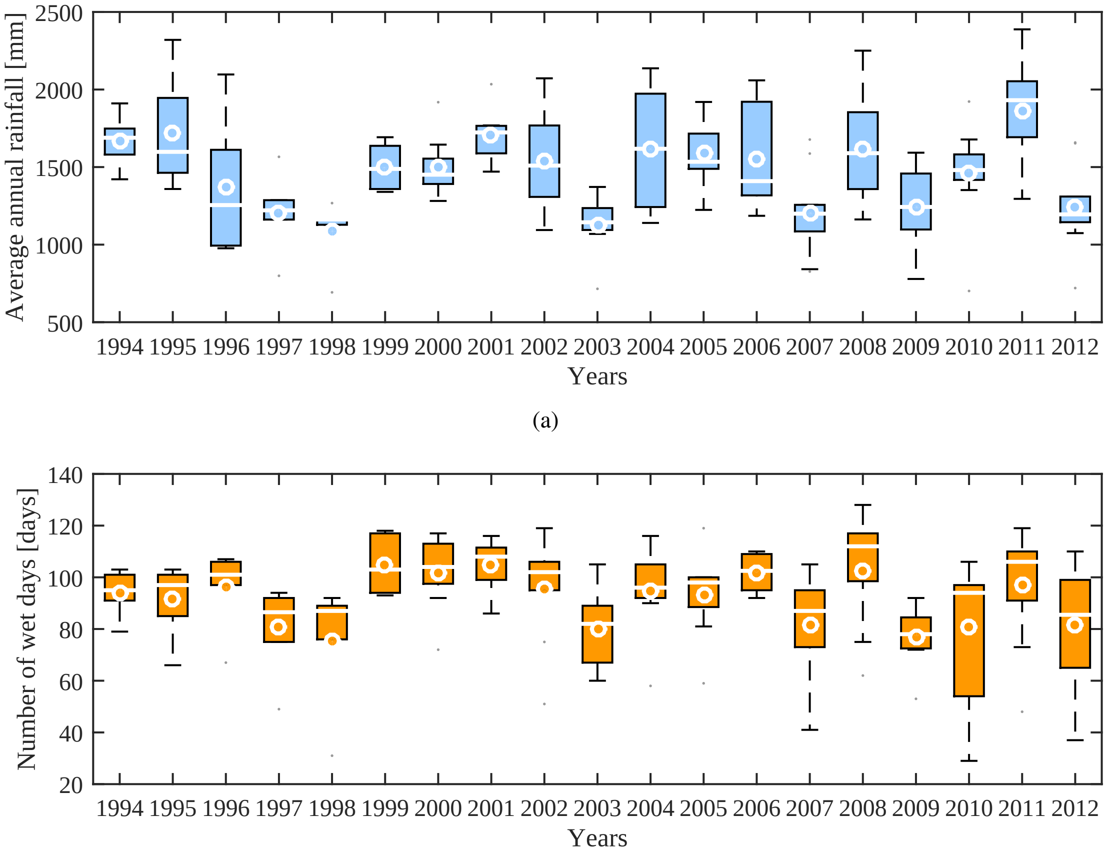


Figure 4.4: Raingauges used for the upper Nan River basin.

Box and whisker plots in Figure 4.5 present annual rainfall information of the upper Nan River basin. The average and median values of the rainfall information are presented by the circle (o) symbol and the line inside the box, respectively. As shown in Figure 4.5a, the annual rainfall of the pilot catchment ranges from 1050 mm to 1850 mm based on the 11 standard raingauges observed during the period 1994–2012. The average annual rainfall is over 1500 mm in the upland region. Moreover, Figure 4.5b provides an information of the spatial rainfall variability in the catchment by using the number of wet days, which was observed from these raingauges, as an indicator. The rainfall threshold of 1 mm was used to define a wet day. In Figure 4.5b, most of the raingauges recorded the annual number of wet days greater than 100 days–for example, in 2006, 2008, and 2011. This indicates large spatial rainfall distributions. Comparing two box plots in Figures 4.5a and 4.5b, most of the raingauges recorded high annual rainfall amounts in these three years. These spatial rainfall distributions and excessive rainfall caused extreme flood events in these years.



(a)

(b)

Figure 4.5: Box and whisker plots for annual rainfall information of the upper Nan River basin based on the 11 raingauges observed during the period 1994–2012: (a) average annual rainfall (b) number of wet days.

Given the exceptional 2011 flood in Nan, the high number of wet days was recorded (as shown in Figure 4.5b). This was consistent with the multiple tropical storms, which moved across northern regions of Thailand in 2011. Tropical monsoon rainfall occurs between June and September accounting for more than 70% of the total annual rainfall. First, during the initial phase of the monsoon season, rainfall transforms dry soil to impervious land. Overlapping of the above-mentioned conditions with additional high-intensity rainfall events extending for a few days induces severe flash flood events, which usually cause damages near the outlet areas of headwater catchments. In the past, flash floods, which occurred in these catchments, often contributed to extreme flood events in the capital of Nan province. These phenomena have caused not only human fatalities but also damaged valuable properties.

4.1.1.2 Radar-rainfall observation

The weather radar station with a 240 km observation range covering the pilot catchment is located at Chiang Rai (19°57’47” N, 99°52’53” E) and is operated by TMD. This radar is a C-band Doppler weather surveillance radar system. It measures radar reflectivity (Z) of a rainfall cloud and hourly stores the radar-based data in form of electronic images. However, these images are available for some heavy rainfall events.

4.1.2 Hydrologic data

The hydrologic data for the pilot catchment were provided by the Royal Irrigation Department (RID). These hydrologic data include water level and discharge data at N.1, N.49, N.64, and N.65. Locations of the monitoring stations in this catchment are shown in Figure 4.4 and Table 4.1. The water level data has been mainly recorded every day and sometimes recorded in hourly time-series during some periods of the wet season. The daily discharge data is converted from measured water levels by using rating curves of each monitoring station developed by RID. The quality of the available rating curves relies solely on polynomial equations that are fitted to measured stage-discharge data. In keeping with Mosley and McKerchar (1993), estimating discharge at stage beyond the measurements by using an extrapolated rating curve developed without consideration of cross-section geometry can result in large estimate errors. Thus, the available rating curves of RID were modified by the author in order to improve these rating curves based on surveyed river cross-sections obtained from RID. The modification approach for extrapolating an rating curve is demonstrated in Section 5.3.

Table 4.1: List of monitoring stations at the upper Nan River basin and their measurements: Q=discharge; WL=water level

No.	Code	Station	River/stream	Latitude [°] E	Longitude [°] N	Drainage area [km^2]	Measurements
1	N.65	Ban Pang Sa	Huai Nam yao	19.23	100.76	637	Q*,WL**
2	N.64	Ban Pha Khwang	Nan	19.01	100.79	3432	Q*,WL**
3	N.49	Ban Nam Yao	Nam Yao	18.99	100.94	155	Q*,WL**
4	N.1	Forestry Office	Nan	18.77	100.78	4609	Q*,WL**

* daily time-series
** hourly records during the wet season

4.2 Upper Ping River basin: an examined region

The upper Ping River basin, an upstream catchment above the P.1 monitoring station at the Nawarat Bridge in Chiang Mai Province, covers an area of 6355 km^2 and is located in the North region of Thailand (Figure 4.6). The river basin extends from the Daen Lao Range in the northern part, with a highest altitude of about 2200 m a.m.s.l., to low-lying areas in the southern part, with a lowest altitude of about 300 m a.m.s.l. (Figure 4.7a). The topography of the river basin is mainly characterized by rugged mountainous terrain, with a mean basin slope of about 34%. The main river, the Ping River, originates at Doi Thuai in the Daen Lao Range and flows southwards passing through the intermontane Chiang Mai region, with a floodplain of about 3 km width. This region is one of the most important cities in Thailand because it has high-population density, economic integration, and historic buildings.

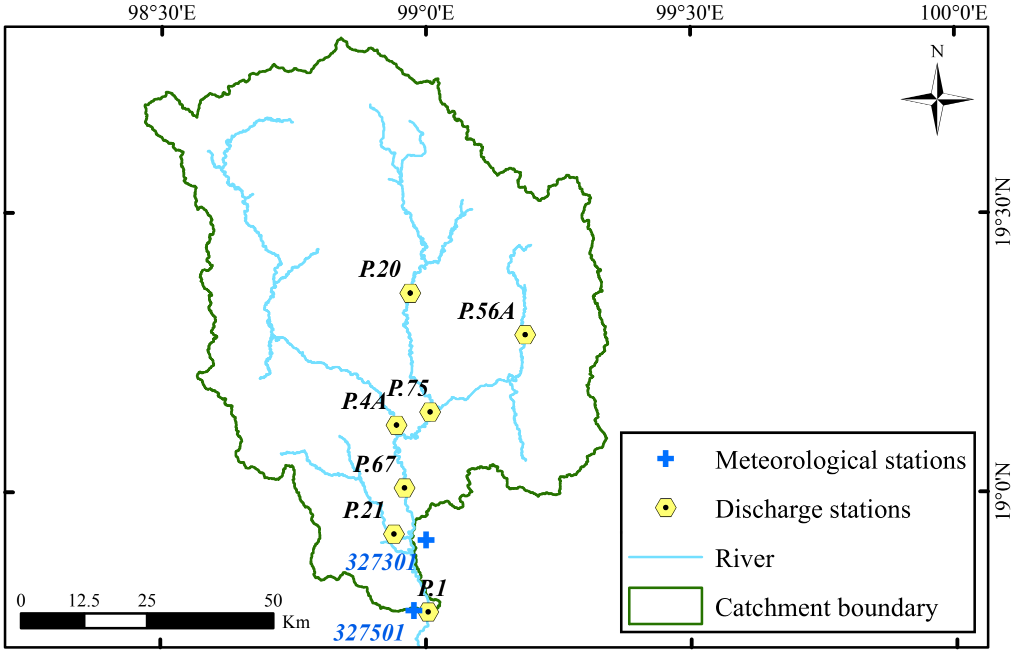


Figure 4.6: Map of the upper Ping River basin.

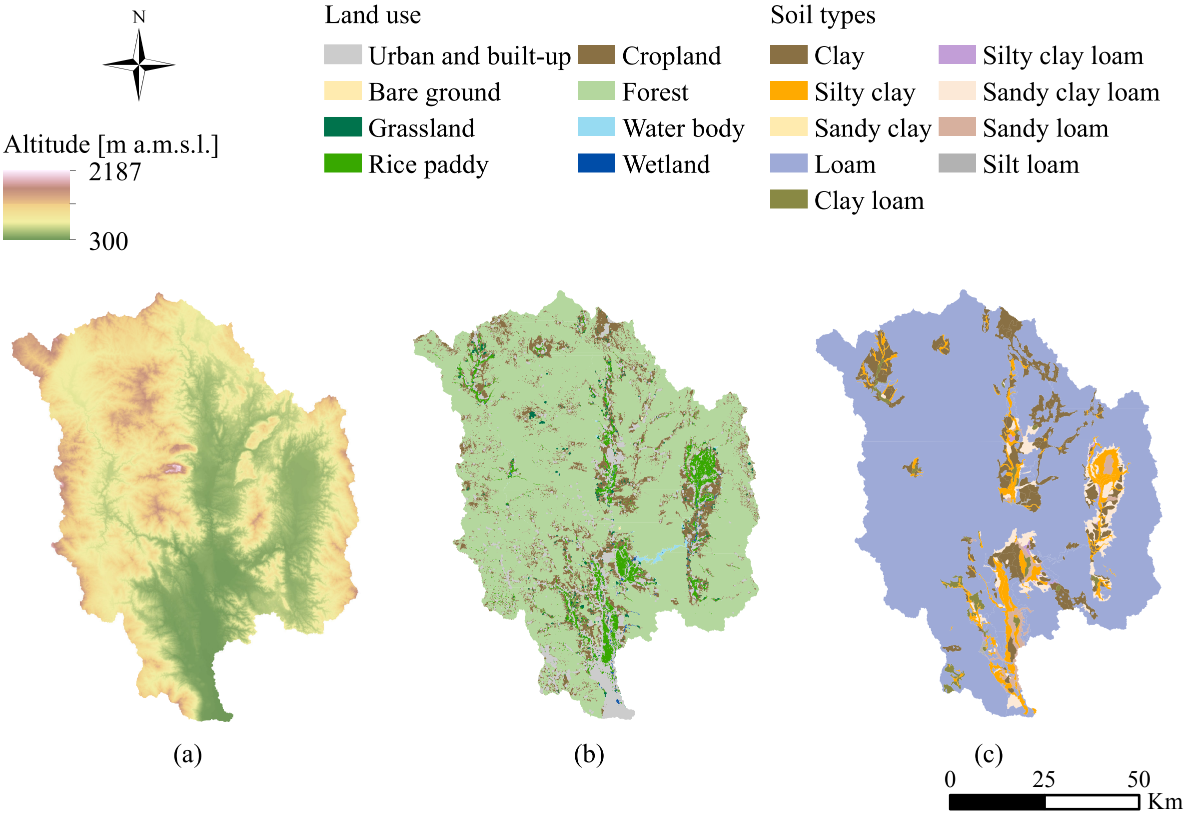


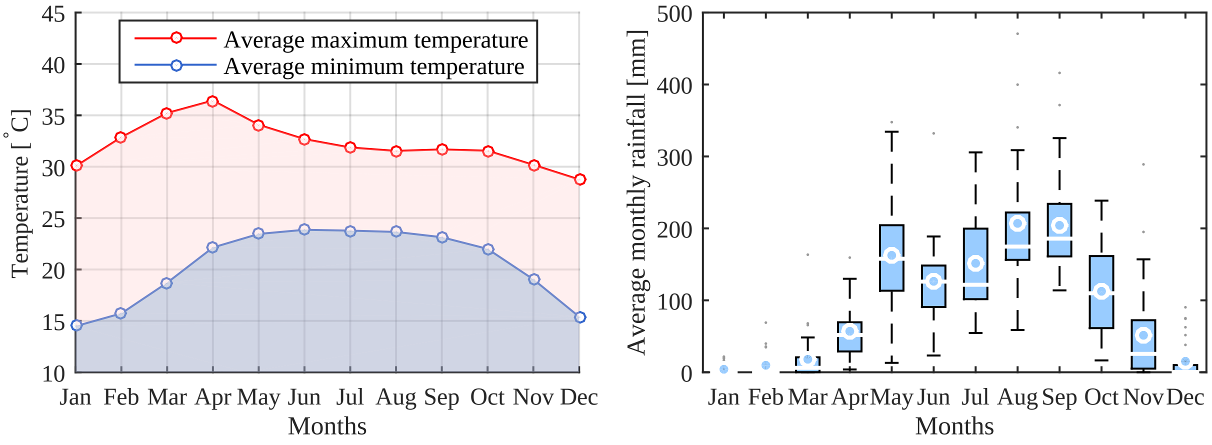
Figure 4.7: The upper Ping River basin: (a) topography of upper Ping River basin (source: RID), (b) land use map based on survey information in 2010 (source: LDD), and (c) soil map (source: LDD).

Land use in the upper Ping River basin can be categorized into seven groups: residential area, bare ground, grassland, rice paddy, cropland, forest, and water body (Figure 4.7b). The predominant land use of the upper Ping River basin was forest, which covered approximately 76% of the total area based on survey information of the Land Development Department (LDD) in 2010. Forested hills are mainly covered by dense deciduous and evergreen forest. According to the LDD, the forest areas have decreased; cropland and residential area, in turn, have increased since the last decades. The cropland and residential area cover about 15% and 4%, respectively, of the basin.

The upper Ping River basin is underlain by folded sediments of Paleozoic sediments interbedded with volcanic rocks, Mesozoic granitic rocks, and Tertiary continental sediments. Upland areas have deep-weathering layers consisting of saprolite varying up to a thickness of ten meters and argillic soil horizons varying up to a thickness of several meters. Surface soils of these areas are predominantly dark brown loam up to a thickness of 25 cm (Wood and Ziegler, 2008). Soils on hillsides are dominated by the deposition of colluvial sediments consisting mainly of sandy clay (Figure 4.7c). In addition, the areas from the foot of the mountains down to the Ping River and its tributaries were formed in water-deposited alluvium (Margane and Tattong, 1999). Most of the soils on these areas are silty clay soils.

4.2.1 Climate

The climate of the upper Ping River basin is influenced by both tropical circulation and annual monsoon. These systems divide the year into two distinct seasons: dry and wet seasons, as mentioned earlier in this Chapter. During the dry season, the average maximum temperature reaches 35°C in March based on two TMD meteorological stations: Mae Jo (327301) and Chiang Mai (327501) stations recording in the period 1981–2010 (Figure 4.8a). In April the average maximum temperature exceeds 36°C over the upper Ping River basin. The average annual temperature is about 26°C. The average maximum and minimum temperatures are 32°C and 20°C, respectively.



(a) (b)

Figure 4.8: Average monthly temperature and rainfall from 30 years of observations in the period 1981–2010 for the upper Ping River basin: (a) plot for average maximum temperature (red solid line) and average minimum temperature (blue solid line) and (b) box and whisker plot for average monthly rainfall.

The wet season usually covers the period of mid-May to October. In this period the upper Ping River basin receives monthly mean rainfall of 100 mm or more. The highest monthly mean rainfall occurs in August about 210 mm (Figure 4.8b). On average rainfall contribution during the wet season accounts for 86% of the total annual rainfall. The average annual rainfall is about 1110 mm based on observations of the TMD meteorological stations between 1981 and 2010.

4.3 Flash flood data

Flash flood information has been published in the reports[2] (in local language) contributed by the Department of Mineral Resources (DMR). The reports present useful information about historic flash flood events and sub-districts affected by these events; however, hydrometric data are not included.

In the absence of reliable runoff data for the sub-basins of the upper Nan River basin and the upper Ping River basin, feasible solutions were based on post-event surveys and remote sensing data. In this dissertation, two field investigations were conducted by the author in order to collect information about historic flash flood events. According to the field investigations, flash flood data were obtained by measurement of high-flood marks from buildings and structures and interviews of eyewitnesses and local people. Figure 4.9 shows high-flood marks from the 2011 flood on some houses in Na Sai village located in the Nam Rim watershed (N1). In addition to point observations of high-flood marks, remote sensing data can be used to estimate flash flood extents (as described in Section 5.5).

Figure 4.9: Na Sai village in the Nam Rim watershed (N1). The horizontal lines indicate the high-flood marks from the 2011 flood. The photographs of W. Lohpaisankrit.

The high-flood marks collected from the field investigations were later used to calibrate a 2D hydrodynamic model (see Section 6.2.2) and to compare with results obtained by the spatial index-based approach for flash flood hazard assessment (see for details Section 7.5).

---

[2] Available on the website: http://www.dmr.go.th/main.php?filename=index (accessed on December 23, 2017)

[This page intentionally left blank]

5 Preliminary Data Analysis

5.1 Rainfall data estimation and selection

5.1.1 Point rainfall observation

The quality of rainfall estimates is a crucial issue for flood forecasting. Conventionally, flood forecasts rely on raingauge networks. The main issue of raingauge networks is that spatial and temporal scales of rainfall data are normally not suitable to fully capture the storm events in small basins. In practice, a spatial distribution of rainfall can be derived by interpolating rainfall data from a raingauge network. Higher raingauge densities may provide higher quality of rainfall estimates and vice versa.

In this dissertation, the rainfall estimates by raingauge measurements were considered as the most reliable source. The available raingauges of the upper Nan River basin is shown in Figure 5.1. Some of the raingauges record rainfall every 24 hours, whereas the others measure rainfall in 15-minute intervals.

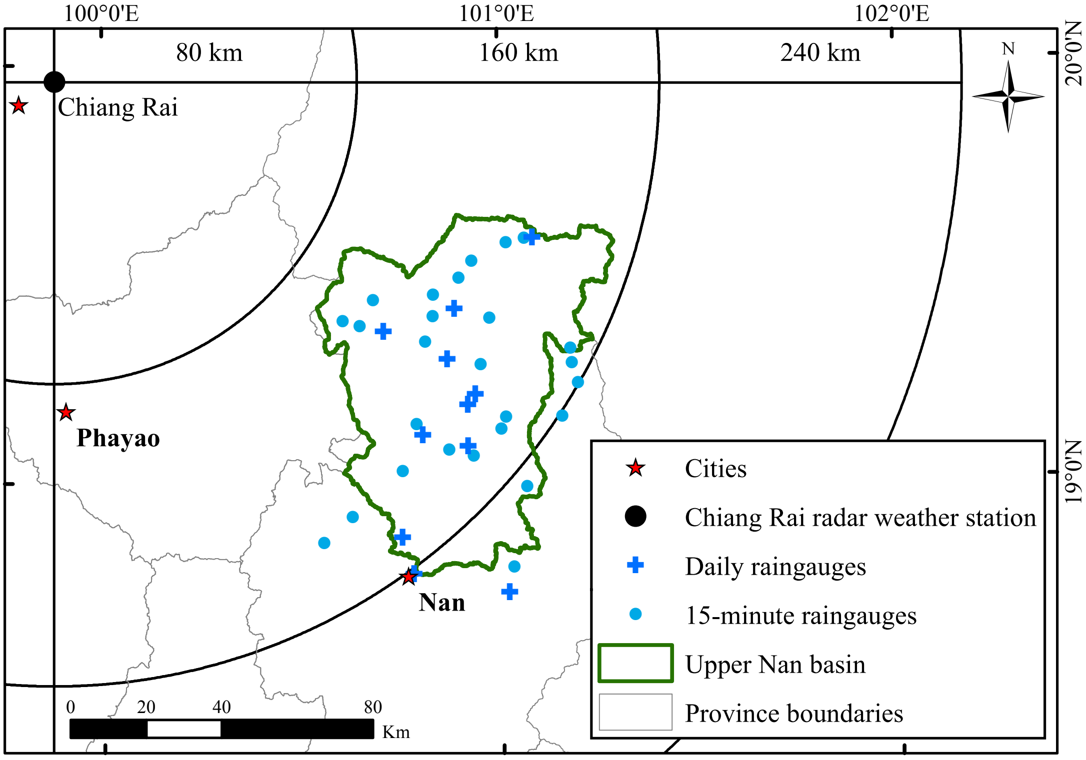


Figure 5.1: Map of the upper Nan River basin, the locations of raingauges, and the location of Chiang Rai radar weather station as well as the 80-, 160-, and 240-km scan radiuses indicated by solid curves.

As previously mentioned, the model-based simulation of flash floods requires high spatial and temporal resolutions of rainfall. The temporal resolution of 15-minute rainfall time-series are sufficient for flash flood simulation. However, the density of the 15-minute raingauge network for the upper Nan River basin was about 380 km^2/station that seemed to be low and not enough for providing high spatial resolutions of rainfall. Moreover, many of the raingauges recorded in15-minute intervals had been operated for only three or four years and contained too few heavy rainfall events for statistical analysis. Therefore, daily rainfall time-series with a long period of about 20 years, were included in the analysis and subdivided by using a disaggregation method (as explained later in Section 5.2). Other alternative sources of rainfall, which provide high resolutions in both spatial and temporal scales, were radar- and satellite-based observations. The observed rainfall data from these sources normally needs to be corrected and compared with observed rainfall data obtained from raingauge measurements before the data were used for further flood analysis.

5.1.2 Radar-rainfall calibration

The C-band radar of the Chiang Rai weather station measures radar reflectivity (Z) of a cloud. The measured reflectivity is expressed in term of dBz (decibel), which is a logarithmic dimensionless unit, and has been stored in the form of radar images. Colors of the radar images, which represent the reflectivity Z, were extracted and decoded using MATLAB tools.

The extreme rainfall event considered in this radar-rainfall correction was the rainstorm between 25–27 June 2011, which caused catastrophic flood events in this respective year. To avoid the bright band effect (Fabry and Zawadzki, 1995) from the radar observation, a network of 21 raingauges, which lay within a 150-km radar coverage range of the Chiang Rai weather station, was accounted for the radar-rainfall analysis. As a result, nine daily raingauges were used as the primary sources for that rainstorm event. The remaining ones were automatic raingauges providing time series at 15-minute intervals. The 15-minute rainfall data were aggregated into hourly time step, since the radar images had hourly temporal and one km^2 spatial resolutions.

To convert the Z values to rainfall intensities (R_i), the climatological $Z - R_i$ relationship was developed through raingauges and radar correction procedures. The Z values of a 25 km^2 grid area covering a raingauge were averaged and matched with observed rainfall data. Afterwards, the mean radar reflectivity was converted to rainfall intensity using the Marshall-Palmer exponential relationship (Marshall and Palmer, 1948) as expressed in Equation (5.1).

$$R_i = \left(\frac{Z}{a}\right)^{-b} \tag{5.1}$$

where a and b are relationship parameters; Z is the reflectivity in dBz; and R_i is the rainfall intensity in $\text{mm}\,\text{h}^{-1}$.

The a and b parameters were calibrated by a trial-and-error procedure. The estimated rainfall intensities from Equation (5.1) were transformed into daily radar-rainfall in millimeter by multiplying with the radar observation interval, which is one hour. The satisfactory values of a and b parameters minimize errors between estimated radar and rainfall observations. Hence, accumulated radar-rainfall during the analyzed rainfall events was statistically compared to accumulated rainfall from raingauges (Figure 5.2). Statistical criteria for the evaluation of the radar-rainfall estimation were mean absolute error (*MAE*), root mean-square error (*RMSE*) and percent bias (*PBIAS*). These criteria are described in Appendix A.2.1.

On the basis of the rainfall data during the extreme rainstorm event of 25–27 June 2011, suitable a and b parameters with 4.57 and 1 were identified, respectively. This b parameter in the $Z - R_i$ relationship was consistent with that of Krajewski and Smith (2002) who noted that it will be close to one for extreme rainfall rates. Figure 5.2 shows a comparison of calibrated radar-rainfall and accumulated rainfall observed by the selected raingauges of the upper Nan River basin during the extreme rainstorm event. The calibrated radar-rainfall underestimated with *MAE* and *RMSE* of 9.9 mm d^{-1} and 12.2 mm d^{-1}, respectively, comparing to the observed rainfall from the gauges. The other statistical criterion, *PBIAS*, indicated overestimation bias of the radar-rainfall with 0.01% comparing to the observed rainfall from the gauges. Moreover, the correlation coefficient between the total amounts of rainfall data recorded by the gauges and the total amounts of raingauge-adjusted radar data during the rainstorm event was 0.73.

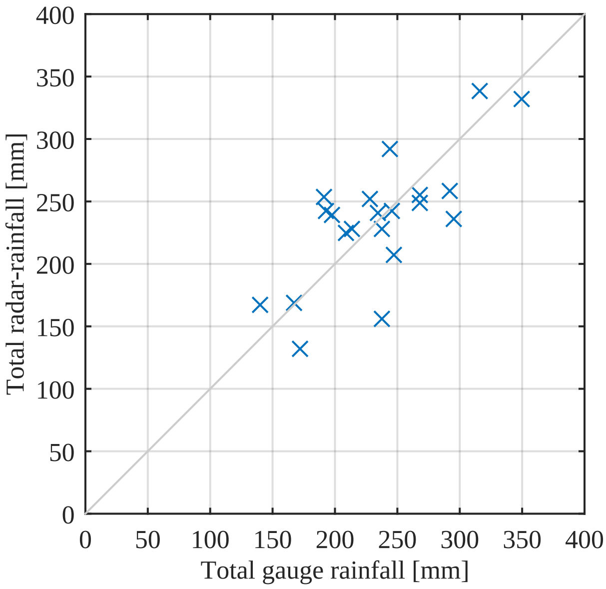


Figure 5.2: Comparison between calibrated radar-rainfall and observed rainfall from raingauges selected for the upper Nan River basin during the extreme rainstorm event of 25–27 June 2011.

5.1.3 Satellite rainfall estimation products

Two satellite rainfall estimation products were analyzed and compared with observed rainfall data from raingauges. The aim of the comparison was to evaluate whether satellite-based rainfall data was suitable for the further analysis of this dissertation. The first product was observed by the Tropical Rainfall Measuring Mission (TRMM[1]) 3B42 version 7. This TRMM estimates a gridded rainfall with a spatial resolution of 0.25 $\times$ 0.25 degrees referring to 25 km length and 25 km width, and a temporal resolution of 3-hour. The other was CMORPH (CPC[2] MORPHing technique) satellite rainfall. Its spatial and temporal resolutions are similar to the TRMM. Figure 5.3 shows the comparison of the two satellite rainfall estimation products, TRMM and CMORPH, and observed rainfall from 25 raingauges selected for the upper Nan River basin during the extreme rainstorm events of 25–27 June 2011.

---

[1] TRMM - a joint mission between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration (JAXA)

[2] CPC - Climate Prediction Center, National Weather Service

According to the comparison results between satellite-based rainfall data and recorded rainfall from the raingauges, satellite-based rainfall data were not consistent with the recorded rainfall data of all raingauges as shown in Figure 5.3. Some raingauges recorded low amount of rainfall, whereas the satellites detect extremely high amount of rainfall. On the other hand, some raingauges delivered very high amounts of rainfall, whereas the satellites detected low amount of rainfall.

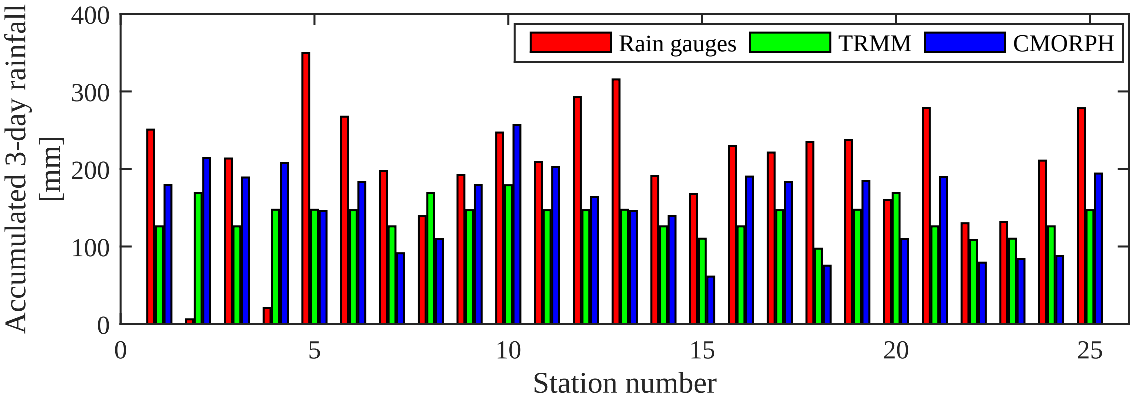
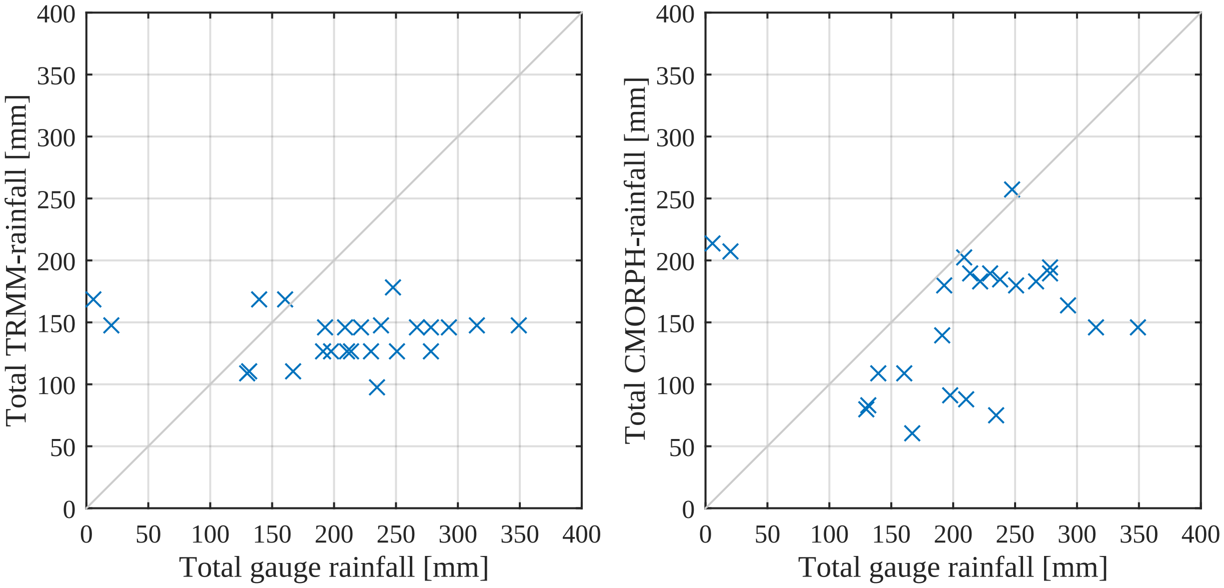


Figure 5.3: Comparison between satellite rainfall estimation products: TRMM (green) and CMORPH (blue) , and observed rainfall from raingauges (red) selected for the upper Nan River basin during the rainstorm event of 25-27 June 2011.

Furthermore, a comparison between the total amounts of TRMM-rainfall data and the total amounts of raingauge data during the rainstorm event is presented in a scatter plot (Figure 5.4a). From the scatter plot, it was found that the correlation coefficient between the rainfall data during the rainstorm event from the two sources was -0.06. Figure 5.4b shows a comparison between the total amounts of CMORPH-rainfall data and the total amounts of rainfall data recorded by raingauges in another scatter plot. It was found that the correlation coefficient between the rainfall data observed by gauges and by CMORPH was 0.09.



(a)

(b)

Figure 5.4: Scatter plots of rainfall data observed by raingauges and by satellite-based observations during the rainstorm event of 25–27 June 2011: (a) comparison between raingauge data and TRMM-rainfall data (b) comparison between raingauge data and CMORPH-rainfall data.

To sum up, among the three remotely rainfall observations (radar and two satellites), the calibrated radar-rainfall data during the rainstorms of 25–27 June 2011 was mostly consistent with the rainfall data observed from the raingauges. Thus, the radar-based rainfall observation was chosen as a supplementary source for the disaggregation of daily rainfall data observed from the raingauges.

5.2 Temporal disaggregation of rainfall

Daily rainfall time-series with a long period of about 20 years are available in the upper Nan River basin. However, the daily resolution is not suitable for characterizing flash floods in small basins (Pekárová et al., 2012), since the short response of flash floods is within minutes to a few hours after the beginning of rain events. Thus, observed daily rainfall data at raingauges in the upper part of the Nan River basin were disaggregated on the basis of hourly rainfall observations. The daily rainfall data were disaggregated into an hourly interval with firstly considering rainfall measured by automatic gauges and then with radar-rainfall observation.

The automatic raingauges operated by the Department of Water Resources (DWR) record rainfall at 15-min intervals and their locations are shown in Figure 5.1. The aforementioned hourly rainfall analysis was the main concern, so the 15-min rainfall data were aggregated to hourly rainfall rates. In order to transfer temporal distribution of hourly to daily time-series, the inverse distance weighting (IDW) method was used to compute weighting factors for hourly rainfall data at neighboring raingauges located within a 10-km radius from a reference daily raingauge. The daily rainfall data at the reference raingauge were disaggregated as a weighted average of the observed hourly rainfall data by using Equation (5.2). As a result, the observed daily rainfall was distributed to follow the mean temporal pattern of hourly rainfall observed at the neighboring raingauges.

$$r^y_{j,k} = r^x_j \left(\frac{R^y_k}{R^x_k} \right) \tag{5.2}$$

where

$r^y_{j,k}$	is the disaggregated hourly rainfall at a reference raingauge [mm];
r^x_j	is the total amount of hourly rainfall observed at all neighboring raingauges in hour *j* (= $\sum_{i=1}^{n} w_i r_i$) [mm];
w_j	is the weighting of each neighboring raingauge (= $d_i^{-2} / \sum_{i=1}^{n} d_i^{-2}$) [-];
r_i	is the total amount of hourly rainfall at a neighboring raingauge *i* [mm];
n	is the number of neighboring raingauges [-];
d_i	is the distance from a reference raingauge to each neighboring raingauge *i* [km];
R^y_k	is the total daily rainfall at a reference raingauge y on day *k* [mm];
R^x_k	is the total daily rainfall at neighboring raingauge(s) on day *k* [mm].

If there was no hourly raingauge within the assigned distance (a 10-km radius) from the daily raingauge, the daily rainfall data were disaggregated with the radar-rainfall observation. Radar-rainfall information was preprocessed and calibrated with observed daily rainfall data (see Section 5.1). The outcome from the preliminary analysis of radar-rainfall observation was a 1-km resolution gridded rainfall data. To extract hourly radar-rainfall data, 5 km x 5 km rainfall grid pixels above a reference raingauge were averaged.

Subsequently, the temporal distributions of hourly rainfall measurements were used to disaggregate daily rainfall data by applying Equation (5.3).

$$r_j^{gauge} = r_j^{radar} \left(\frac{R_k^{gauge}}{R_k^{radar}} \right) \tag{5.3}$$

where

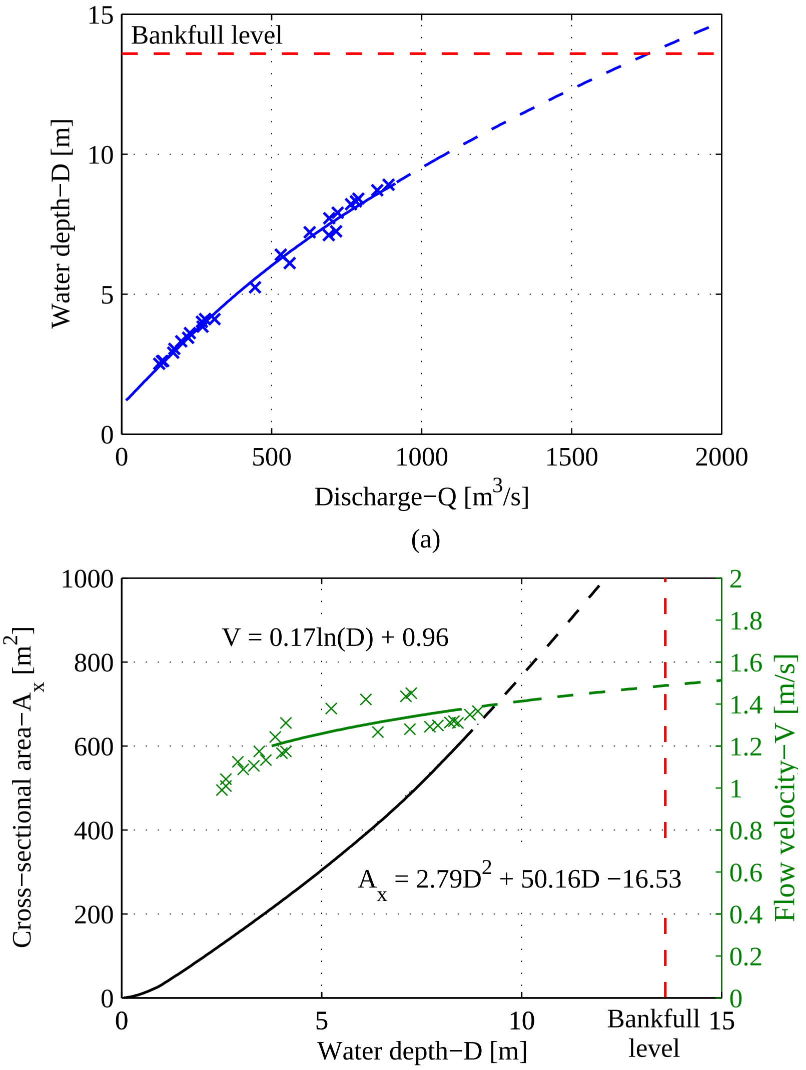
r_j^{gauge} is the disaggregated hourly rainfall at a reference raingauge in hour *j* [mm];
r_j^{radar} is the average gridded radar-rainfall in hour *j* [mm];
R_k^{gauge} is the total daily rainfall at a reference raingauge on day *k* [mm];
R_k^{radar} is the total amount of hourly gridded radar-rainfall on day *k* [mm];

5.3 Improvement of rating curves

In hydrology, a rating curve is a graphical plot of a stage-discharge relationship at a cross section of a river, usually at gauging stations. In Thailand, the Royal Irrigation Department (RID) operates main river flow gauges across the country. The RID normally reports daily mean discharge in an official database, although the flow depth measurements have been recorded on an hourly basis at almost all gauging stations in the wet season. Actually, river flow depths at each gauging station have been converted to river flow by using satisfactory rating curves.

The establishment of a reliable rating curve is a difficult task because direct measurements of river discharge are expensive, time-consuming, and sometimes not performable in extreme flow conditions such as floods and low flows. As a result, the upper and lower parts of the rating curve are usually missing or highly uncertain. To estimate the discharge at stages beyond the measurements, several techniques have been applied to fit measured stage-discharge data and extrapolated from the measurements to cover the entire range of flows. The RID has represented the stage-discharge relationships based on power or polynomial equations. According to Mosley and McKerchar (1993), the extrapolation of a rating curve without consideration of cross-section geometry can result in large estimate errors. Therefore, the development of the rating curves in this dissertation included the information of surveyed river cross-sections.

Since river cross profiles have changed over time, the measurements of stages and discharges between 2008 and 2011 were considered in the rating curve improvement at the N.1, N.64 and N.65 gauging stations. For example, the measured cross-sectional area of the N.64 station and its depth stages were related and used to develop a cross-sectional area function corresponding to water depth as shown in Figure 5.5.



(a)

(b)

Figure 5.5: Extrapolation of a rating curve at N.64 monitoring station (a) Original rating curve extrapolated with the dash blue line (b) Functions to be used for extrapolation when D > 9.0 m: A_x = f(D), black; V = f(D), green.

The original rating curve was improved and extended through the velocity-area method (Mosley and McKerchar, 1993; Shaw et al., 2011) . On the basis of the velocity-area method, discharge can be determined with the following equation:

$$Q = A_x V \tag{5.4}$$

where

Q	is discharge [$m^3 s^{-1}$];
A_x	is the cross-sectional area of the river [m^2];
V	is the average velocity [$m s^{-1}$];

In addition, flow velocity was obtained by using the Equation (5.4) and the velocity function corresponding to water depth was developed. These two functions: cross-sectional area and velocity functions were used for extrapolating high discharges where stages were higher than nine meters as an example of the N.64 rating curve shown in Figure 5.5.

5.4 Selection of sub-basins for flash flood analysis

As flash floods usually occur in steep and small basins, this research work focused on the mountainous and headwater watersheds of the upper Nan River basin. The criteria for selecting sub-basins were that:

(i) the size of sub-basins was smaller than about 350 km^2 because basins with drainage areas less than 350 km^2 have a high potential for flash flood occurrence (Marchi et al., 2010);

(ii) time of concentration of the sub-basins was shorter than six hours (see Section 3.1);

(iii) these sub-basins were suffered from flash floods in history.

The five sub-basins that drain into the Nan River as shown in Figure 5.6 were selected for flash flood analysis. The first sub-basin was the Nam Rim watershed (N1) located in the East of the upper Nan River basin. The Nam Rim watershed covers an area of 76 km^2. The second sub-basin was the Nam Hui watershed (N2) located in the Northwest of its parent basin. The Nam Hui watershed covers an area of 43 km^2. The third sub-basin was the Nam Lae watershed (N3) located in the North of its parent basin. The Nam Lae watershed covers an area of 44 km^2. The fourth sub-basin was the Nam Pua watershed (N4) located in the East of its parent basin. The Nam Pua watershed covers an area of 196 km^2. The last sub-basin was the Nam Khwang watershed (N5) located in the East of its parent basin. The Nam Khwang watershed covers an area of 155 km^2.

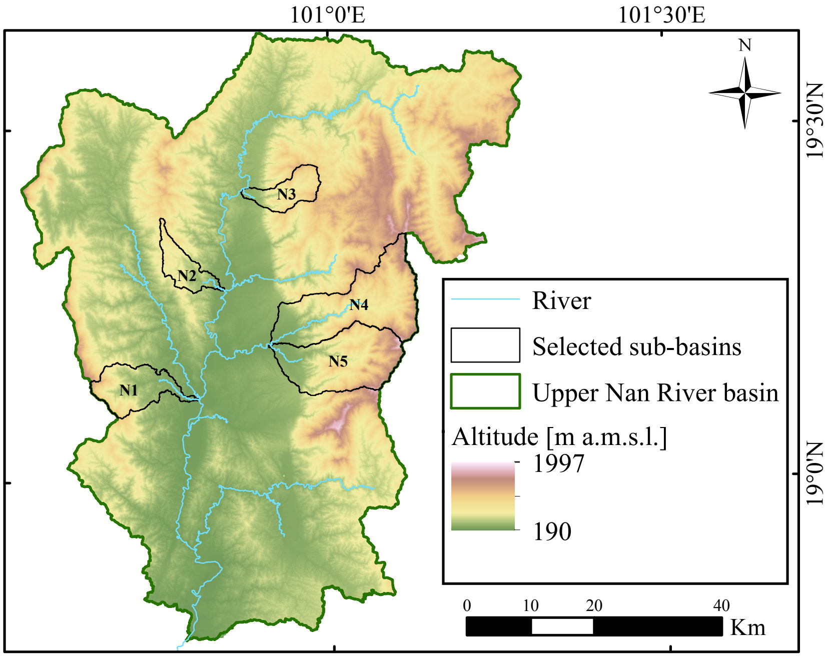


Figure 5.6: Selected sub-basins in the upper Nan River basin: Nam Rim watershed (N1), Nam Hui watershed (N2), Nam Lae watershed (N3), Nam Pua watershed (N4), and Nam Khwang watershed (N5).

5.4.1 Analysis of basin characteristics

Morphometric characteristics of the five sub-basins were extracted by using ArcMap[3] 10.3 from a 10-m digital elevation model (DEM). These morphometric characteristics of the sub-basins can be represented by morphometric parameters listed in Table 5.1. The morphometric parameters can be categorized into two aspects: areal and linear aspects.

Table 5.1: Morphometric parameters of the selected sub-basins in the upper Nan River basin

No.	Morphometric parameters	Unit	Sub-basin				
			N1	N2	N3	N4	N5
1	Basin area (A)	km^2	76.3	42.8	44.1	196.0	155.0
2	Basin perimeter (P_b)	km	68.5	54.8	46.9	107.7	76.7
3	Basin length (L_b)	km	17.2	15.0	11.2	27.5	20.4
5	Mean basin slope (S_m)	%	37.6	34.7	42.0	45.4	46.3
4	Basin relief ratio (S_r)	m/m	0.047	0.028	0.055	0.035	0.051
6	Basin relief (H_b)	m	1232.7	804.7	1041.0	1691.6	1688.7
7	Main channel length (L_m)	km	18.0	11.8	10.5	29.5	24.6
8	Mean bifurcation ratio (Rb_{mean})	[-]	4.8	3.6	4.6	3.7	4.4
9	Mean length ratio (Rl_{mean})	[-]	1.7	1.4	1.8	2.1	1.6
10	Length of longest flow path (L_{lfp})	km	26.0	27.5	16.6	47.2	32.6

The areal aspects analyzed here were basin area (A), basin perimeter (P_b), basin length (L_b), basin relief (H_b), mean basin slope (S_m), and basin relief ratio (S_r). The first four areal aspects of a given basin can be depicted in Figure 5.7. The estimation of the last two areal aspects is described later in this section. These morphometric parameters were selected because they have been considered to be important factors, which affect the hydrologic behavior of a basin (Youssef et al., 2011). In addition, the morphometric parameters were used to estimate basin concentration time, which is an important indicator of the hydrologic response to intense rainfall (WMO, 1994). The estimation of basin concentration time is presented later in Section 5.4.2.

The selected sub-basins of the upper Nan River basin are smaller 350 km^2. Among the selected sub-basins, the Nam Hui watershed (N2) was the smallest basin with an area 43 km^2 and its perimeter was 55 km. On the contrary, the Nam Pua watershed (N4) was the largest basin with an area of 196 km^2 and its perimeter was 108 km. The basin length of the all sub-basins ranged from 11 to 28 km. The mean basin slope (S_m) parameter was computed by using spatial analyst tools in the ArcMap. As can be seen in Table 5.1, this S_m parameter of all the sub-basins was relatively high, which ranged from 35% to 46%. Another morphometric parameter was basin relief, which is the elevation difference between the basin divide and outlet (Strahler, 1952). This parameter controls stream gradient, which thereby influences flood patterns (Gallagher, 1999). Among the sub-basins, the Nam Hui watershed (N2) had a least basin relief of 805 m, which was considered as a high relief drainage basin. As stated by Patton and Baker (1976), a basin with high relief may manifest a high flood potential even given low rainfall amounts of rainfall. Moreover, basin relief ratio (S_r) defined by Schumm (1956) as the ratio between basin relief (H_b) and basin length (L_b)

---

[3] ArcMap is an application used in ArcGIS software, which is a geographic information system developed by the Environmental Systems Research Institute (ESRI)

was relatively high for all the sub-basins. The Nam Hui watershed (N2) had a relief ratio of 0.028, which was the lowest relief ratio among all the sub-basins, whereas the Nam Lae watershed (N3) had the highest relief ratio of 0.055. According to Ghoneim et al. (2002), these high values of basin relief and relief ratio indicate short time of runoff concentration.

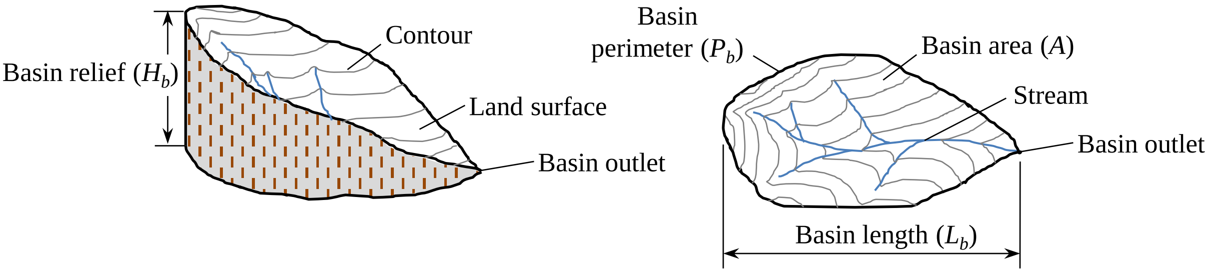


Figure 5.7: Areal aspects of a basin (taken from Ritter (1978), modified).

In this dissertation, the linear aspects of the selected sub-basins such as main channel length (L_m), mean bifurcation ratio (Rb_{mean}), mean length ratio (Rl_{mean}), and length of longest flow path (L_{lfp}) were analyzed. The main channel of a basin was computed on the basis of Horton (1945) that defined the highest order stream as the main channel. The Nam Hui watershed (N2) had the shortest main stream of 12 km, whereas the Nam Pua watershed (N4) had the longest main stream of 30 km. In addition to the analysis of the main stream, its drainage network was analyzed and ordered by using the stream order analysis of Strahler (1957). As a result, the bifurcation ratio and length ratio for all the sub-basins were obtained. The mean bifurcation ratio ranged from 3.6 to 4.8 and the length ratio ranged from 1.4 to 2.1. According to Ritter (1978), the values of bifurcation ratio usually range between 3.0 and 5.0 for watersheds formed by homogeneous geologic structures that do not distort their drainage patterns. The last linear aspect that relates to concentration time of basins was length of the longest flow path. The longest flow path of the sub-basins lain between 17 and 47 km.

The analysis of the aforementioned morphometric characteristics revealed that the selected sub-basins have a high potential for flash flood occurrence. Moreover, these morphometric characteristics were considered as important parameters for estimating concentration time of basins as discussed in the following section.

5.4.2 Estimation of basin concentration time

Five standard methods selected by the author for estimating basin concentration time (t_c) were Giandotti (1934), US Corps Engineers (1936), California Culverts Practice (1942), Johnstone and Cross (1949), and Dooge (1956) methods as summarized in Table 5.2. These five formulas are empirical relations of selected basin parameters such as basin drainage area, length of main channel, length of the longest flow path, elevation difference between the mean altitude of the basin and outlet, elevation difference between basin divide and outlet and average basin slope. These five formulas to estimate concentration time of the sub-basins were chosen because they were developed from small to midsize rural and mountainous basins. In addition, their required basin parameters could be derived from the available digital elevation model.

Table 5.2: Summary of basin concentration time formulas

Methods	Formulas for t_c [hours]	Remarks
Giandotti (1934)	$t_c = \dfrac{4\sqrt{A} + 1.5L_m}{0.8\sqrt{H_m}}$	This formula was developed from basins in central and northern Italy (170-70 000 km^2) (Radice et al., 2012).
US Corps Engineers (1936)	$t_c = \dfrac{0.191L_m^{0.76}}{S_r^{0.19}}$	This formula was developed from mountainous basins in California (de Almeida et al., 2014).
California Culverts Practice (1942)	$t_c = 0.95\left(\dfrac{L_{lfp}^3}{H_b}\right)^{0.385}$	This formula was developed from small mountainous basins in California (Chow et al., 1988; Grimaldi et al., 2012).
Johnstone and Cross (1949)	$t_c = \dfrac{0.4623\sqrt{L_m}}{S_r^{0.25}}$	This formula was developed for basins with drainage areas between 64 and 4200 km^2 (Li and Chibber, 2008; Grimaldi et al., 2012).
Dooge (1956)	$t_c = \dfrac{0.365A^{0.41}}{S_r^{0.17}}$	This formula was derived from 10 rural basins ranging from 145 - 948 km^2 in Ireland (de Almeida et al., 2014).

Note: A = basin area [km^2]; L_m = length of main channel [km]; L_{lfp} = length of the longest flow path [km]; H_m = the elevation difference between the mean altitude of the basin and outlet [m]; H_b = the elevation difference between the basin divide and outlet [m]; S_r = basin relief ratio or average basin slope [m/m].

Results of concentration time computed by using the above-mentioned five formulas for all the sub-basins are shown in a box-and-whisker plot (Figure 5.8). In this plot, the horizontal axis is named by the five sub-basins and the vertical axis represents the time of concentration in hour. The whiskers, which are two lines outside the boxes, extend to the maximum and minimum values of the concentration time. Segments inside the boxes represent the median and points show the average values of the concentration time for the five sub-basins. As it can be observed, the average values of the concentration time for all the sub-basins are shorter than six hours.

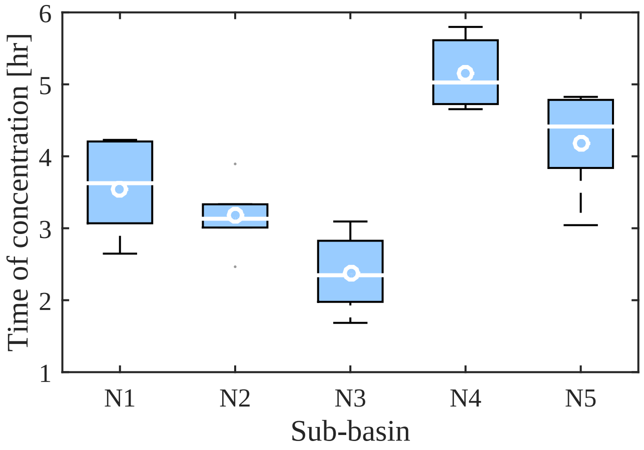


Figure 5.8: Box-and-whisker plot of concentration time for the selected five sub-basins of the upper Nan River basin: Nam Rim watershed (N1), Nam Hui watershed (N2), Nam Lae watershed (N3), Nam Pua watershed (N4), and Nam Khwang watershed (N5).

5.5 Post-flash flood survey

5.5.1 Field investigation

The purposes of the field investigations by the author were to gather information about historic flood events from local eyewitnesses and inhabitants. Witness interviews were carried out in order to understand local flood characteristics, and timing and duration of the flood events. Local flood characteristics were such as whether water flow or debris flow and presence of woody debris in a stream. Another useful information gained from witnesses was flash flood occurrence. Before flash flood occurrence in the study regions, there were usually prolonged rainfall in a few days. Consequently, this had increased moisture in the soil. The soil moisture conditions and hilly characteristics of catchments often cause flash floods when the catchment areas receive more moderate or heavy rainfall. Sometimes, rain falls over upper parts of a catchment and flash floods occur in lower areas where rain did not even fall. In addition, local eyewitnesses and inhabitants provided indirect flood estimation; they provided the timing of floods and identified where high-flood levels were marked. High-flood marks from buildings and structures were measured by a measuring tape, and located by a GPS receiver. These past flood locations were transformed into a digital format by the ArcMap software and were later used to compare with results of further analysis in this dissertation.

The author conducted two field investigations about historic flash floods in the five sub-basins of the upper Nan River basin as shown in Figure 5.9a and the four sub-basins of the upper Ping River basin as shown in Figure 5.9b, respectively. For the five sub-basins of the upper Nan River basin, they are described in Section 5.4. For the four sub-basins of the upper Ping River basin, they are described as follows.

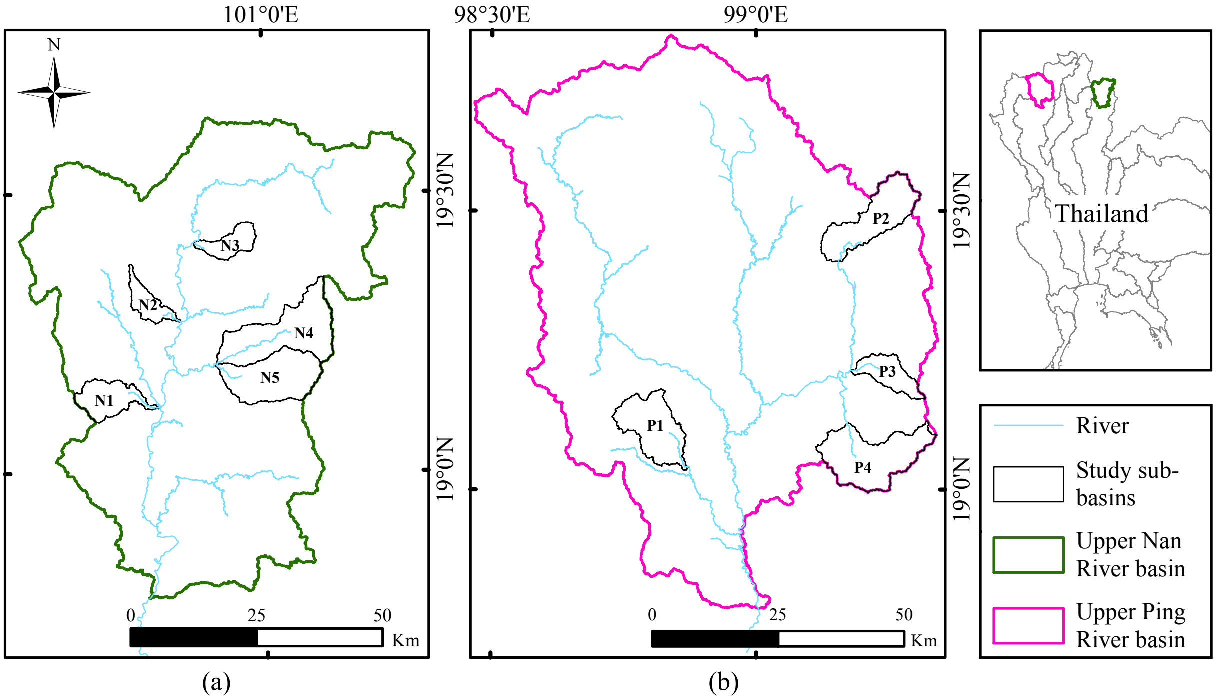


Figure 5.9: Study sub-basins and their locations (a) five sub-basins of the upper Nan River basin and (b) four sub-basins of the upper Ping River basin.

The first sub-basin was the Nam Mae Hao watershed (P1) located in the Southwest of the upper Ping River basin. The Nam Mae Hao watershed covers an area of 129 km^2. The second sub-basin was the Nam Mae Ngat watershed (P2) located in the Northeast of its parent basin. The Nam Mae Ngat watershed covers an area of 139 km^2. The third sub-basin was the Nam Mae Ngat branch watershed (P3) located in the East of its parent basin. The Nam Mae Ngat branch watershed covers an area of 60 km^2. The last sub-basin was the Nam Mae Khot watershed (P4) located in the Southeast of its parent basin. The Nam Mae Khot watershed covers an area of 181 km^2. These four sub-basins of the upper Ping River basin including the five sub-basins of the upper Nan River basin are summarized in Table 5.3.

Table 5.3: Study sub-basins located in the upper Nan River basin and the upper Ping River basin

Code	Name	Area [km^2]	Main river basin
N1	Nam Rim watershed	76.3	Upper Nan River basin
N2	Nam Hui watershed	42.8	
N3	Nam Lae watershed	44.1	
N4	Nam Pua watershed	196.0	
N5	Nam Khwang watershed	155.0	
P1	Nam Mae Hao watershed	129.2	Upper Ping River basin
P2	Nam Mae Ngat watershed	139.4	
P3	Nam Mae Ngat branch watershed	60.3	
P4	Nam Mae Khot watershed	181.1	

5.5.2 Remote sensing of flood extent

Field measurements can provide high-flood marks, but no continuous flood boundaries. It may be costly and time-consuming to conduct spatially detailed observations of flood locations through the field measurements. Therefore, remote sensors on satellites can provide data to supplement the field measurements.

In this dissertation, inundated areas were extracted from satellite imagery of the Landsat-7. The Landsat-7 operated and maintained by the United States Geological Survey (USGS) was launched by the National Aeronautics and Space Administration (NASA) in 1999. The Landsat-7 has gathered images of all landmass and near-coastal areas on the Earth in order to observe changes on the land surface (Arvidson et al., 2001). It carries the Landsat Enhanced Thematic Mapper Plus (ETM+) sensor, which has acquired the images in eight spectral bands. The images of seven spectral bands (bands 1 to 7) have a spatial resolution of 30 m. The resolution of images acquired by the panchromatic band 8 is 15 m. Landsat-7 images are available to download free of charge on several websites (e.g. https://earthexplorer.usgs.gov).

The procedure of a remote sensing technique for extracting inundated areas consisted of four steps as presented in Figure 5.10. The first step was a composition of three Landsat-7 spectral bands. These spectral bands were band 4, band 5, and band 3 chosen because their combinations can differentiate the water bodies from other land uses. For example, band 4, band 5, and band 3 images of the Landsat-7 observation over the Nam Mae Hao (P1) on 26 July 2011 are shown in Figure 5.11. The band 4, band 5, and band 3 images were composited in a Red, Green, Blue (RGB) composition in order to visualize the imagery in color as displayed in Figure 5.12a. As it can be seen in this figure, flooded areas appear very dark gray[4]. In addition, it was important to note that clouds and their shadows can cause errors in a classification procedure because the shadow is captured in dark color, which is representing wet areas. Thus, areas covered by clouds and their shadows must be corrected in the last step.

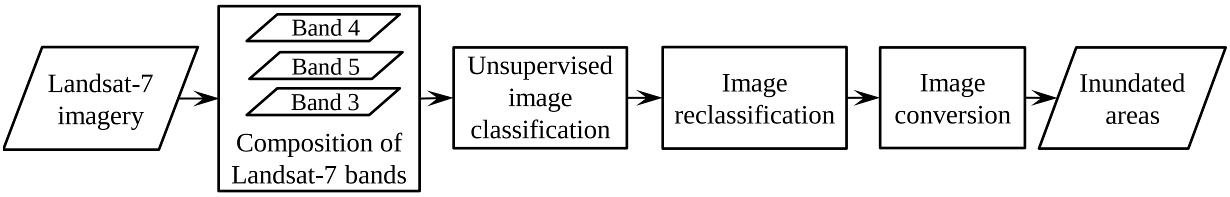
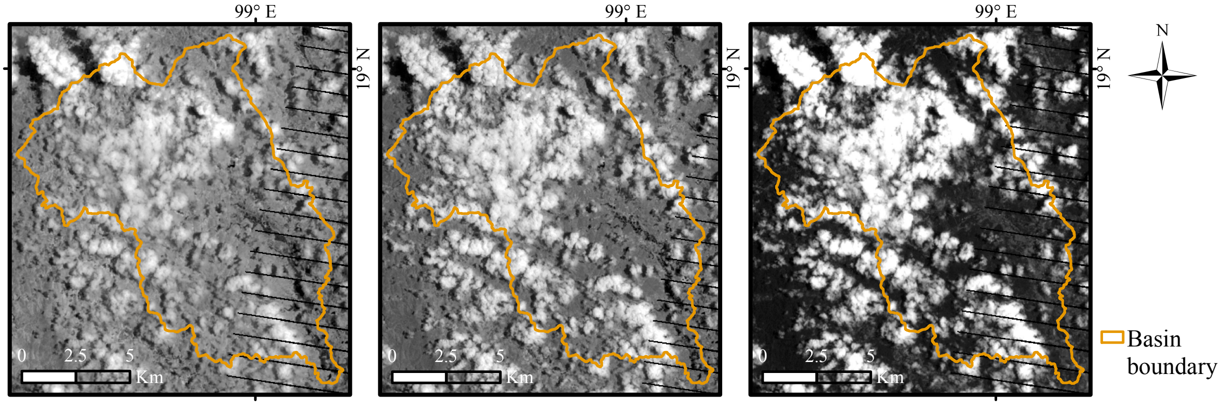


Figure 5.10: A flow chart of inundated area extraction (in ArcMap software).



(a) (b) (c)

Figure 5.11: Satellite images observed by Landsat-7 on 26 July 2011 over the Nam Mae Hao (P1) watershed detected by (a) band 4, (b) band 5, and (c) band 3.

The second step was an unsupervised image classification. Images of RGB band combinations were classified by using ISO[5] cluster unsupervised classification algorithms (Tou and Gonzalez, 1981) implemented in the ArcMap. Outcomes obtained from the unsupervised image classification were digital images stored in raster graphics, which are made of generally rectangular pixels. A large number of unknown pixels was classified and divided into a number of classes. Each class was given by a value. As a result, the spectral classed images represented by a number of values having the same number of the classes were created.

The third step was an image reclassification. The classes that result from the unsupervised image classification were based on natural groupings of image values. Therefore, it was necessary to reclassify the spectral classed images. In order to identify the classes of these images, ares were compared with reference data such as maps and information obtained from field investigations. Afterwards, the images were reclassified into two classes: wet and dry pixels.

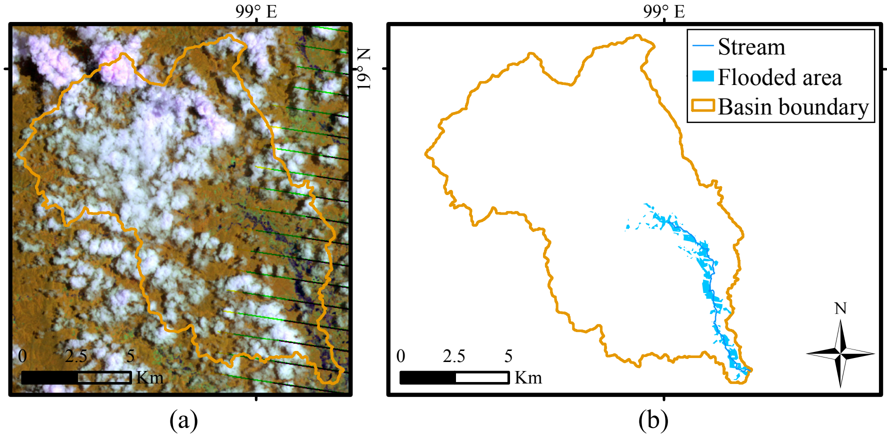
The final step was an image conversion. As previously mentioned, areas covered by clouds and their shadows can cause errors in delineating inundated boundaries. Therefore, these areas were identified as dry pixels. The wet pixels were transformed into polygons that were used to represent inundated areas. For

---

[4] For interpretation of color in Figure 5.12a, the reader is referred to an electronic version of this dissertation

[5] ISO stands for Iterative Self Organizing.

example, inundated areas extracted from the Landsat-7 imagery observed on 26 July 2011 over the Nam Mae Hao watershed (P1) is shown in Figure 5.12b.



(a) (b)

Figure 5.12: Composition of Landsat-7 satellite bands observed on 26 July 2011 over the Nam Mae Hao watershed (P1): (a) Results of a RGB band combination (Red: band 4, Green: band 5, Blue: band 3) and (b) flood areas extracted from the RGB band combination.

The procedure for extracting inundated areas of the Landsat-7 imagery was applied to the remaining study watersheds of the upper Ping River basin – namely, the Nam Mae Ngat watershed (P2), the Nam Mae Ngat branch watershed (P3), and the Nam Mae Khot waterhsed (P4). The results of extracted inundated areas were later used for comparison with results obtained by the spatial index-based approach for flash flood potential assessment (see for details Section 7.5).

[This page intentionally left blank]

6 Approach

This chapter describes the approaches used in this dissertation to carry on flash flood potential assessment in tropical regions. To assess flash flood potential in space, two different approaches were developed and applied: the integrated modeling approach and the spatial index-based approach outlined in Figure 6.1. This figure is self-explaining. Generally, the integrated modeling approach can provide results with high accuracy. However, it normally requires the collection and management of numerous detailed data, which are expensive and time-consuming processes. In many cases, such data are not available. On the other hand, the spatial index-based approach, which was initiated here, required digital elevation data only.

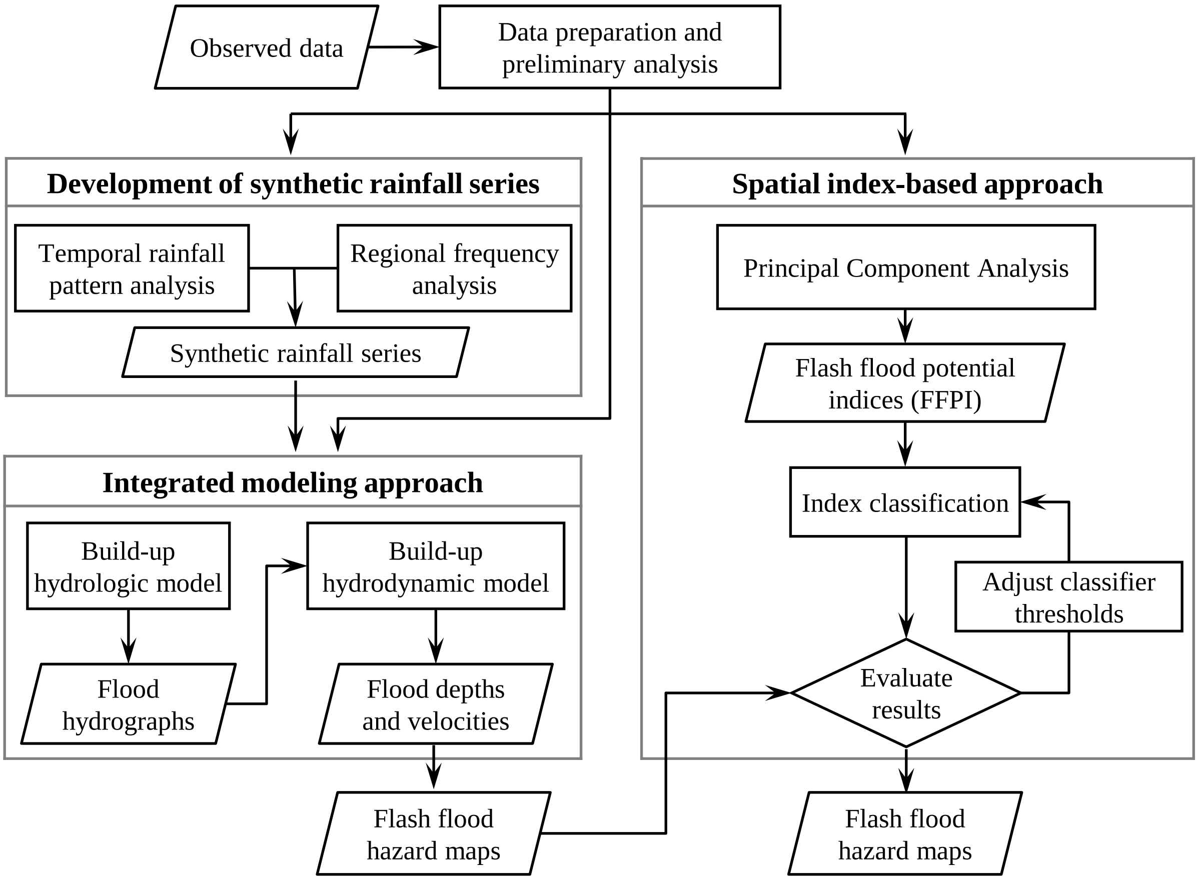


Figure 6.1: Overview of the approaches.

The new spatial index-based approach was the main innovative contribution of this dissertation. For developing the spatial index-based approach, the upper Nan River basin was selected as a pilot catchment. To validate the new approach, the methodology was applied to analyze and identify flash flood hazard areas in the upper Ping River basin (see Chapter 4).

To analyze flash floods in the study region, it required temporal and spatial data. The temporal data consisted of the hydrological and meteorological data. The spatial data included topography, land use and soil maps. Rainfall observations were considered in both temporal and spatial distributions because they play a crucial role in flash flood occurrence. The preparation and preliminary analysis of the data have been described in Chapter 5.

As it has been discussed, one of the challenges is to deal with the lack of observed runoff data. To estimate flash flood runoff in the ungauged basins, it was necessary to synthesize rainfall series. Procedures for developing the synthetic rainfall series are described in Section 6.1. These synthetic rainfall series were used as an input to the first approach for delineating flash flood hazard areas under different rainfall scenarios.

In short, two main research approaches were applied to assess flash flood hazard in this dissertation (Figure 6.2). The first approach created flash flood hazard maps with the help of the integrated hydrologic and hydrodynamic models. In this approach, the hydrologic model (PANTA RHEI) was applied to the upper Nan River basin and the hydrodynamic model (MEADFLOW) was applied to the five watersheds located in the basin (as shown in Section 5.4). The second approach, which was a spatial index-based approach, derived flash flood potential indices (FFPI) to determine areas prone to flash floods. This second approach was applied to the five watersheds of the upper Nan River basin and the four watersheds of the upper Ping River basin (as shown in Section 5.5). This approach was elaborated on the basis of the principal component analysis, a multivariate analysis technique. Afterwards, the FFPI were classified and then used to map flash flood hazards.

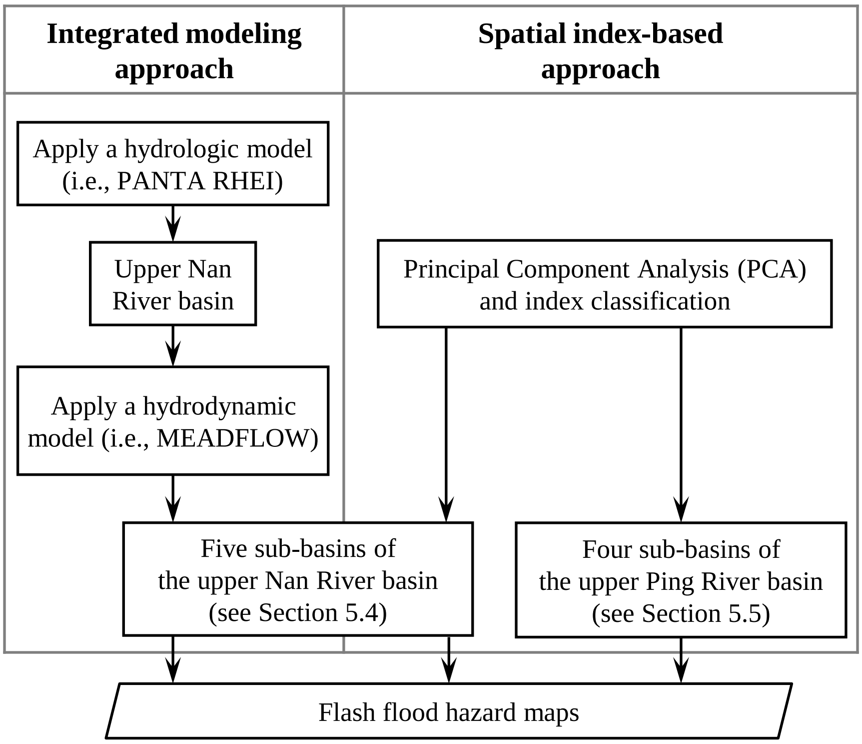


Figure 6.2: Two research approaches selected for flash flood hazard mapping and application areas.

For the upper Nan River basin, the results obtained by means of the spatial index-based approach were compared with the ones of the integrated modeling approach. Moreover, the distance between the two approaches will be analyzed and discussed for further improvement. For the upper Ping River basin, the results obtained by the spatial index-based approach were compared with flood inundation extents extracted from the Landsat-7 imagery (see Section 5.5). These results are discussed in Chapter 7.

6.1 Development of synthetic rainfall series

Owing to the shortage of highly resolved spatio-temporal rainfall fields for analyzing flash floods in the selected sub-catchments, it was necessary to develop synthetic rainfall series from the available daily rainfall data. The development of synthetic rainfall series consisted of two main procedures: temporal rainfall pattern and regional frequency analysis.

6.1.1 Analysis of temporal rainfall patterns

The analysis of temporal rainfall patterns was based on radar-rainfall observations. These observations were available in forms of radar-rainfall images at hourly intervals. The temporal patterns of rainfall were characterized by catchment-averaged rainfall series (Yucel, 2015). Afterwards, these patterns were averaged and then transformed into a dimensionless mass curve. The dimensionless mass curve, which represents the temporal rainfall pattern of the pilot catchment, is displayed in Section 7.1.

6.1.2 Regional frequency analysis

The regional frequency analysis (RFA) has become necessary when high-resolution rainfall time series are too short by statistical standards. Therefore, annual maximum 24 h (AM24H) rainfall data observed by raingauges for 16 years in average were chosen to analyze the rainfall frequency characteristics of the region (the upper Nan River basin as shown in Figure 4.4). RFA can be summarized into four procedures: (i) screening of the data, (ii) testing of regional homogeneity, (iii) identifying of a regional frequency distribution, and (iv) estimating of regional frequency relationships. These procedures are drawn in a schematic diagram as shown in Figure 6.3.

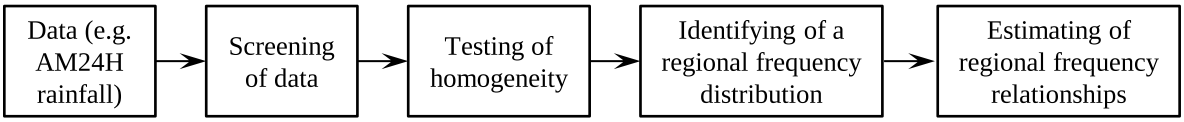


Figure 6.3: Schematic diagram of regional frequency analysis of data (i.e., annual maximum 24 h (AM24H) rainfall data).

The purpose of the data screening was to exclude outliers and inconsistencies in data in order to maintain the statistical properties of the data. The data screening procedure consisted of two steps. The first step individually examined outliers of the AM24H rainfall data at each station. To screen the outliers, Grubb's test (Grubbs, 1969) was applied. The outliers that depart significantly from the remaining data were excluded from the analysis because they may distort the statistical properties of data (Chow et al., 1988). Afterwards, raingauges were considered together as a whole in order to examine their inconsistencies in the group of the raingauges. The inconsistencies were analyzed by using a discordancy measure (D_i) based on *L*-moment techniques. This discordancy measure compares between the *L*-moment ratios of a station (i.e., a raingauge) and the mean *L*-moment ratios of a group of stations. If the D_i value of a station is large, the station is determined to be discordant. According to Hosking and Wallis (1997), the D_i value of a station is considered as "large" when the D_i value is greater than the critical value given in Table 6.1. Therefore, any station with a large value of D_i was discarded. Afterwards, the remaining raingauge stations were selected for the regional homogeneity test.

Table 6.1: Critical values for the discordancy measure (Hosking and Wallis, 1997)

Number of stations in region	Critical value	Number of stations in region	Critical value	Number of stations in region	Critical value
5	1.333	9	2.329	13	2.869
6	1.648	10	2.491	14	2.971
7	1.917	11	2.632	≥15	3
8	2.140	12	2.757		

To test the homogeneity of the selected raingauges, a heterogeneity measure (*H*) based on cluster analysis was applied. This heterogeneity measure compares the dispersion of *L*-moment ratios for a group of stations with the expected dispersion of these in the case of a homogeneous region. The expected dispersion of regional *L*-moment ratios was established by performing Monte Carlo simulations of a homogeneous region with artificial sites. The number of artificial sites were assumed to be equal to the number of the observed stations. Moreover, the artificial sites were assumed to have the same record lengths as the observed data of the stations. For example, a large number (greater than or equal to 500) of Monte Carlo simulations generate samples govern by a four-parameter kappa distribution. Each simulation produces the number of the samples being equal to the number of the artificial sites. For each site, RFA computes *L*-moment ratios such as *L*-coefficient of variation (*L*-Cv), *L*-coefficient of skewness (*L*-Cs), and *L*-coefficient of kurtosis (*L*-Ck) from its samples. These *L*-moment ratios for all the sites were averaged according to their record lengths and then used to estimate the dispersion of the regional *L*-moment ratios. Furthermore, RFA computed the dispersion of *L*-moment ratios from the group of the observed raingauge stations. Afterwards, the dispersions of the *L*-moment ratios calculated from the simulated and observed data series were evaluated by using the heterogeneity measure.

In order to determine whether a region would be homogeneous or not, Hosking and Wallis (1993) suggested that a region can be considered as 'acceptably homogeneous' if $H < 1$, 'possibly heterogeneous' if $1 \leq H < 2$, and 'definitely heterogeneous' if $H \geq 2$. This heterogeneity measure in this dissertation was calculated by applying an nsRFA[1] package developed by Viglione (2014) in R, an open source programming language and software environment (R Core Team, 2015).

The regional distribution function was chosen on the basis of the Z-statistic (Z^{dist}). According to Hosking and Wallis (1997), the Z^{dist} is a goodness-of-fit measure, which indicates how well *L*-kurtosis of theoretical distributions fits with the regional mean *L*-kurtosis of the observed data. The theoretical distribution with the minimum absolute Z^{dist} value was the most preferable.

Once the most suitable frequency distribution was identified for the region, its regional distribution pa-

---

[1] nsRFA stands for non-supervised Regional Frequency Analysis

rameters were used to estimate regional frequency relationships that were depth-duration-frequency (DDF) curves. As the regional frequency analysis was based on the 24 h duration of the observed rainfall data in this dissertation, the shorter durations were necessary to be used to calculate the DDF curves. Therefore, proportions of 24 h rainfall values for shorter durations were assumed as Bell/SCS type II ratios (Bell, 1969; Awadallah et al., 2011). These rainfall proportions and regional distribution parameters were used to compute the DDF curves for all the daily raingauges.

As a result of the previously described procedures, the combinations of the temporal rainfall pattern and the DDF curves were synthetic rainfall series. These rainfall series were a key input for hydrological modeling in order to estimate potential flash floods. Details about the obtained series are described in Section 7.1.

6.2 Integrated hydrologic and hydrodynamic modeling of flash floods

The observed hydrological and meteorological time-series data, and spatial data of basin characteristic information were used to build up an integrated hydrologic and hydrodynamic modeling system for the five selected sub-basins of the upper Nan River basin as previously shown in Figure 5.6. After this system was calibrated and validated, the synthetic rainfall were used as input and flash flood hazard maps with several rainfall scenarios were created. Figure 6.4 shows a framework of the integrated hydrologic and hydrodynamic models for flash flood hazard mapping. Detailed procedures of the framework are explained as following sections.

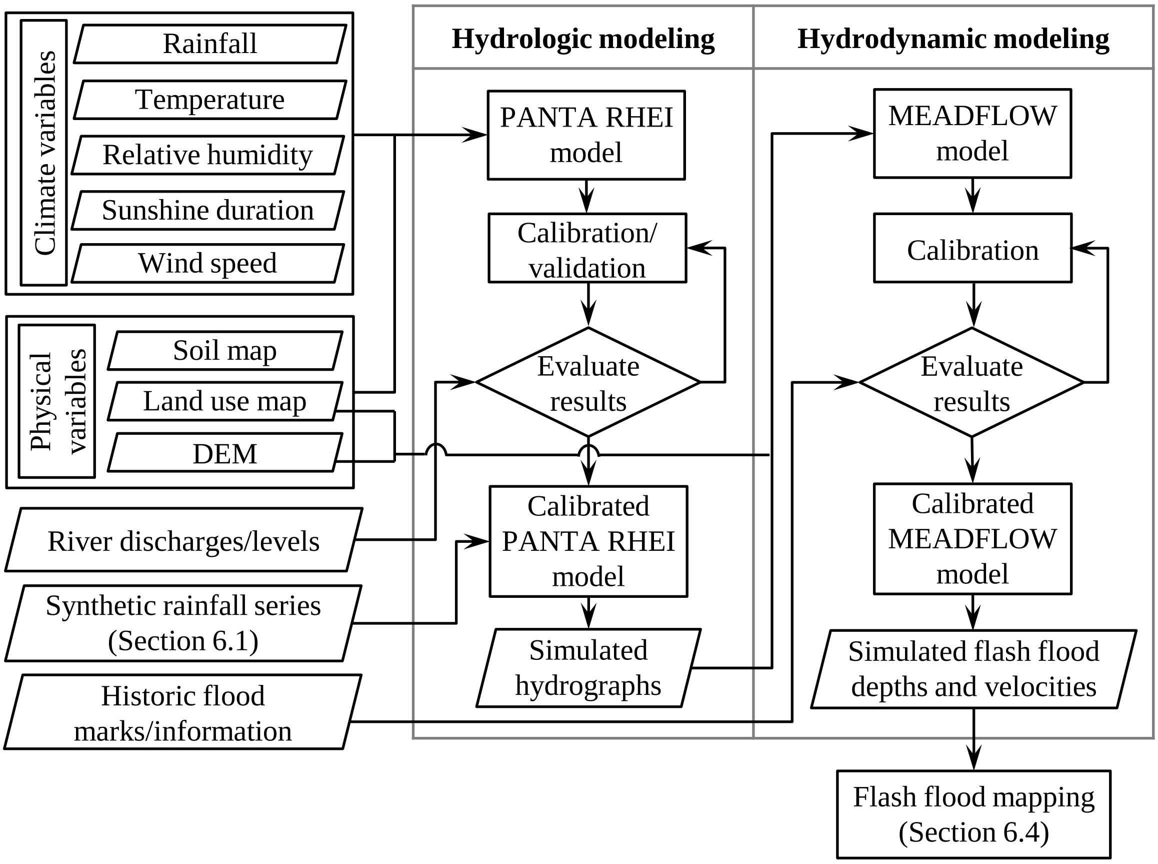


Figure 6.4: Framework of integrated hydrologic and hydrodynamic models for flash flood hazard mapping.

6.2.1 PANTA RHEI hydrologic modeling

6.2.1.1 Description of the hydrologic modeling system PANTA RHEI

PANTA RHEI is a semi-distributed and physically based hydrologic modeling system, which has been developed by the Leichweiss Institute for Hydraulic Engineering and Water Resources (LWI-HYWAG), University of Braunschweig, and the Institut für Wassermanagement IfW Incorp., Braunschweig, Germany (LWI-HYWAG and IFW, 2015). It has been used in various scientific and practical applications at national and international levels. There are three principal versions of PANTA RHEI. First, the scientific version of PANTA RHEI has been used in both water quantity studies (e.g., Kreye, 2015) and water quality studies (e.g., Lorenz et al., 2017). Second, the operational version of PANTA RHEI has been applied by the Flood Early Warning Center of Lower Saxony in Germany for real-time flood forecasting (e.g., Meon et al., 2015; Riedel et al., 2017). Third, the planning version of PANTA RHEI has been used for climate impact studies on the water balance in the Lower Saxony region (Hölscher et al., 2012). Moreover, PANTA RHEI was applied in international research projects using an extended ecohydrologic version to study river water quality as a part of the integrated project "Environmental and Water Protection Technologies of Coastal Zones in Vietnam (EWATEC-COAST)" (Meon and Phuoc, 2014; Lorenz et al., 2017).

PANTA RHEI has been designed with a modular structure which allows users to choose hydrologic approaches based on their research objectives and available data. The temporal discretization of PANTA RHEI calculation is adaptive and the common time step is one hour. The spatial discretization can be categorized into three levels: gauged watersheds, sub-basins and hydrologic response units (HRU) in the mesoscale. The watersheds are divided into sub-basins, which contain the HRU. A hydrologic response unit is an area with a unique land use and soil mapping combination. This HRU represents hydrologic characteristics of an area. Normally, a number of HRU are aggregated into sub-basins. This spatial discretization is also called semi-distributed.

6.2.1.2 Application of PANTA RHEI

The objective of applying PANTA RHEI in this dissertation was to estimate runoff in ungauged basins located in the upper part of the Nan River basin. This basin was discretized into a network of 554 sub-basins, with an average area of 8 km^2. Every sub-basin contained at least one HRU. In total, there were 3927 HRU inside the 554 sub-basins, which were defined according to different combinations of land use and soil types. In addition, the sub-basins also contained the information of topographic characteristics – namely, surface slope, stream slope, stream length. The topographic characteristics were derived from a digital elevation model (DEM) with a 10-m resolution obtained from the Royal Irrigation Department by using the watershed delineation techniques (e.g., Maidment, 2002).

After hydrologic and topographic characteristics of the sub-basins were derived from physical variables (DEM, land use and soil types), the PANTA RHEI model required climatic variables for building it up as shown in Figure 6.4. The climatic variables were rainfall, temperature, relative humidity, sunshine duration, and wind speed. As previously mentioned in Section 4.1.1, the observed climatic variables were mainly available in daily time series, excepting rainfall data. The rainfall data were available in daily and sub-daily time series.

In order to improve the model performance, daily rainfall data were disaggregated with hourly rainfall data observed at neighboring stations and radar-rainfall images. The method of temporal rainfall disaggregation is described in Section 5.2.

The PANTA RHEI model was set up for the upper Nan River basin covering an area of 4609 km^2 and run at hourly time steps. This model consisted of three main modules representing the hydrologic processes of a river basin as shown in Figure 6.5. The first module was runoff formation, which included evapotranspiration, interception, and direct runoff processes. The potential evapotranspiration was estimated by using the Penman-Monteith approach (Monteith, 1973). The interception loss was estimated according to Rutter et al. (1971). The direct runoff was computed on the basis of the SCS-CN method. However, the values of CN in PANTA RHEI were timely modified in order to vary the CN values based on soil moisture conditions throughout the entire simulation process (Riedel, 2004). This modified SCS-CN method estimated infiltration and percolation in terms of proportions to runoff. The second module was runoff concentration, which conveys rainfall excess to the watercourse. The runoff concentration was calculated by using the linear reservoir functions of four discharge components such as surface flow, interflow, groundwater flow, and baseflow (Chow et al., 1988). The last module was runoff routing, which composed of linear storage functions and translation components (LWI-HYWAG and IFW, 2015).

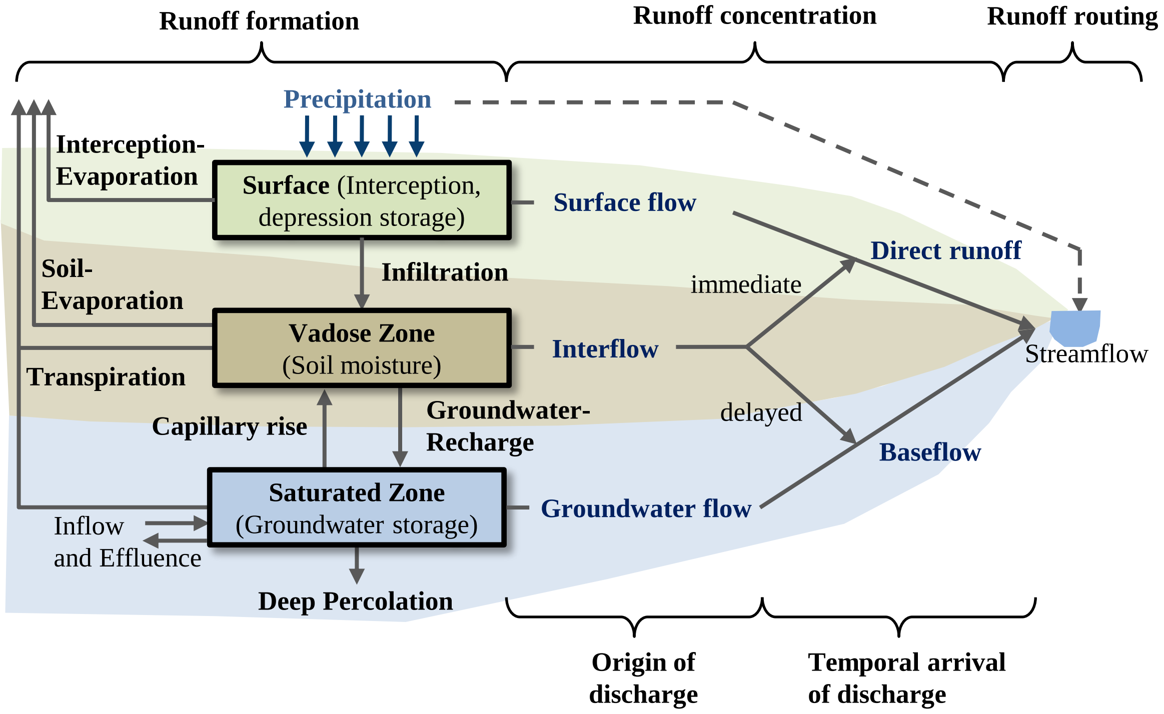


Figure 6.5: Simplified catchment longitudinal section and schematic representation of main hydrologic processes in PANTA RHEI (modified from Baumgartner and Liebscher, 1996).

Once the PANTA RHEI model was constructed, two extreme flood events of June 2011 and August 2008 at the Nan River basin were used to calibrate and validate this model. In the calibration simulated and observed hydrographs were statistically compared at the three river discharge stations: N.65, N.64 and N.1. The locations of these stations are displayed in Figure 4.4. The other station (N.49) has recorded river discharge information for its upstream drainage area of 155 km^2. There was only one raingauge station, which observed rainfall between 2010 and 2012, in this area. Owing to a low density of raingauges, it may cause overparameterization in the model calibration. To avoid this, the observed discharge time series of the N.49 station were included to simulate hydrographs in the PANTA RHEI model. The calibration period was between 25 and 30 June 2011. The model parameters were calibrated by using an optimization algorithm (Gelleszun et al., 2015) for multi-objective functions with efficiency index (*EI*) and relative peak error (*RPE*) statistics described in Appendix A.2.1. Furthermore, the PANTA RHEI model was validated by comparing simulated hydrographs with observed hydrographs between 17 July and 11 August 2008. After the model was calibrated and validated, the calibrated parameters were carried out to simulate flash flood hydrographs for sub-basins of the upper Nan River basin. Results of the calibration, validation, and synthetic simulation of flash floods are presented in Section 7.2.

6.2.1.3 Sensitivity analysis of model parameters

Five sub-basins of the upper Nan River basin were selected for studying flash floods. Once hydrologic behavior of the upper Nan River basin was successfully represented by the PANTA RHEI model, the calibrated model parameters of the parent basin were transferred to its sub-basins. As mentioned earlier, the five sub-basins were ungauged basins. Therefore, it was necessary to assess the effect of model parameters on outcomes of the model through sensitivity analysis.

For the sensitivity analysis of the model parameters, six model parameters from the three components of the runoff generation were selected. Three of the parameters were maximum interception storage, a factor of potential evapotranspiration, and a factor of percolation in the runoff formation component. Two of the parameters were a factor of interflow storage and a factor of groundwater storage in the runoff concentration component. The last parameter that was a flow velocity factor was selected from the runoff routing component.

To generate various values of the model parameters for the sensitivity analysis, the Latin hypercube sampling technique introduced by McKay et al. (1979) was applied. This Latin hypercube sampling technique randomly generated different values of all the model parameters ranging within boundaries that were set to be between -20% and +20% of the calibrated model parameters. As a result, a large number of the combined parameter sets was obtained. Subsequently, these combined parameter sets were used by the PANTA RHEI model to simulate hydrographs.

6.2.1.4 Estimation of design flash floods

For the estimation of design flash floods, two assumptions were made. The first assumption was that the initial soil moisture content was assumed to be at the field capacity for simulating design flash floods. The other assumption was that a set of calibrated model parameters obtained from the parent basin (i.e., the upper Nan River basin) can simulate flash floods in its sub-basins because the set of calibrated model parameters was able to reproduce extreme floods in the river floodplains of the parent basin. This assumption was supported by morphological characteristics of the parent basin. Mean bifurcation ratio of the upstream catchment above the N.64 gauging station (the parent basin of the five selected sub-basins as shown in Section 5.4) was 3.8 meaning that the catchment was formed by homogeneous geologic structures (Ritter, 1978).

Two types of design rainfall distributions were used for estimating design flash floods in the five study sub-basins. The first type of design rainfall distribution was temporal and spatial variations of extreme rainfall. The temporal variations of extreme rainfall, so-called synthetic rainfall series, were developed by combining regional temporal rainfall pattern and frequency analysis described in Section 6.1. To simulate the spatial variation of extreme rainfall, the synthetic rainfall series of raingauges were spatially interpolated by the inverse distance weighting method implemented in the PANTA RHEI model. The purpose of using these temporal and spatial variations of rainfall based on regional frequency analysis of extreme rainfall events was to design regional rainfall scenarios for all the selected five sub-basins. Afterwards, the PANTA RHEI model transferred the temporal and spatial variations of extreme rainfall inputs to their corresponding flash flood hydrographs for the five sub-basins.

The second type of design rainfall distribution was a spatially uniform rainfall distribution. The purpose of using a spatially uniform rainfall distribution was to examine various combinations of rainfall amount and duration that may lead to flash flooding in an individual sub-basin. To estimate design rainfall amounts, it was necessary to analyze historical rainfall with certain durations in order to understand rainfall characteristics of the region that may cause flash floods. It has been noted that flash floods in small to midsize watersheds can be triggered by intense rain falling within particular durations (Marchi et al., 2010). Very intense rainfall within a short duration may not always cause a flash flood in such watersheds. In the upper Nan River basin, maximum rainfall data with durations of 1, 3, 6, 12, and 24 hours are available at the Muang Nan meteorological station (331201). These rainfall data were used to simulate flash flood hydrographs by the PANTA RHEI model.

In this dissertation, the duration of six-hour rainfall was considered as the critical storm duration because the duration is longer than the basin concentration time of each selected watershed. On the basis of the well-known Rational method (see Section 3.2.2), the critical storm duration can produce the peak discharge for small basins. This critical storm duration was normally determined to be equal or greater than the basin time of concentration (Ponce, 1989).

6.2.2 MEADFLOW hydrodynamic modeling

6.2.2.1 Description of the hydrodynamic modeling system MEADFLOW

MEADFLOW is a 2D hydrodynamic modeling system and was developed by the Ingenieure für Wasser, Umwelt und Datenverarbeitung (IWUD) GmbH. It focuses on the identification of spatial flow distribution in rivers, floodplains and wetlands. This hydrodynamic modeling system has been applied in many practical purposes. For example, MEADFLOW was combined with a groundwater model in order to simulate the interaction of surface water and groundwater in a floodplain (Koch et al., 2004). Another application of the hydrodynamic modeling system is the simulation of inundated areas due to consequences of dam failure (Anhalt and Meon, 2008).

MEADFLOW is implemented in Gecko-2D, a graphical user interface software package. This software package includes pre- and post-processing modules, which can be used to create a triangular irregular network (TIN) from topographic information (Leismann and Meon, 2002). Other fundamental information such as flow resistances and hydraulic boundary conditions is required for spatial flow simulation. Moreover, MEADFLOW uses the complete shallow-water equations (see Section 3.3.2). The equations are solved by the numeric algorithms based on the Galerkin finite element method with triangular elements

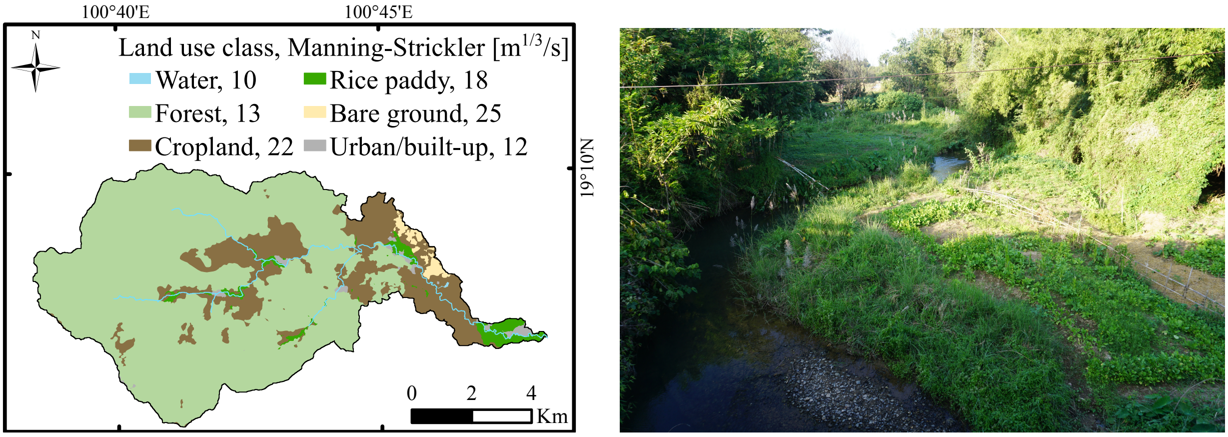
(Zienkiewicz et al., 2014). If compared with other 2D model software, MEADFLOW has low computation times (Leismann and Meon, 2002).

6.2.2.2 Application of MEADFLOW

MEADFLOW was applied to route peak discharges simulated from PANTA RHEI and to simulate spatial flood depths and velocities at downstream areas of the selected sub-basins in the upper Nan River basin. The reasons why MEADFLOW was selected for hydrodynamic modeling were its short computational times and user-friendly output options. To build up the MEADFLOW model, it required model boundaries and parameterization, which were topography, roughness parameters, and boundary conditions.

On the subject of topography, MEADFLOW generates a TIN for representing flood terrain from a digital elevation model. The DEM data with a 10-m resolution used in this dissertation were transformed into three-dimensional (3D) coordinates, which represent longitude, latitude, and altitude of the DEM data, in the grid data module of MEADFLOW. These 3D coordinates were linked in the mesh generation module of MEADFLOW and then a TIN that represented topographic areas of a sub-basin was created.

The Manning-Strickler (k_{St}) roughness coefficients were used to calibrate the MEADFLOW model. The k_{St} roughness coefficients are the reciprocal of Manning's n roughness coefficients. These roughness coefficients depend on land use units in terms of their roughness. For example, land use of the Nam Rim watershed was classified into six classes: water, forest, cropland, rice paddy, bare ground, and built-up units. Figure 6.6a. shows the land use units of the Nam Rim watershed and their corresponding k_{St} roughness coefficients, which were obtained from the model calibration. The estimation of the roughness coefficients based on land use can be found in the literature (e.g., Chow, 1959; DVWK, 1991).



(a)

(b)

Figure 6.6: Land use map of Nam Rim watershed: (a) Land use classes and Manning-Strickler (k_{St}) and (b) photograph of the Nam Rim stream and vegetation, facing downstream (taken by W. Lohpaisankrit).

In addition to the topography and roughness parameters, two boundary conditions necessary for starting the hydrodynamic computation were the upstream and downstream boundaries. The upstream boundary was assigned by using simulated peak floods from the PANTA RHEI hydrologic model. Meanwhile, the downstream boundary was assumed by a bankfull elevation of the stream. This approach for assigning the upstream and downstream boundaries was applied because there was no observed runoff and stream level in the selected sub-basins.

MEADFLOW was run in the steady state mode. To clarify, the upstream boundary condition was given as a peak of discharge and the upstream boundary condition was set as a bankfull elevation throughout the simulation. The purpose of using this mode was to save computational times as this dissertation paid attention to maximum flood extends. To calibrate the MEADFLOW model, the k_{St} roughness coefficients were adjusted by using a trial-and-error approach and simulated flood depths were compared with maximum flood marks, which occurred in 2011. These maximum flood marks were obtained from field investigations and eyewitness reports. Figure 6.6a presents the values of the k_{St} roughness coefficients for their corresponding land use classes, which were obtained from the calibration. Figure 6.6b shows the Nam Rim stream and its densely vegetated floodplains. The k_{St} roughness coefficient of $10\,\mathrm{m}^{1/3}\,\mathrm{s}^{-1}$ was used for the water body. For forest, cropland, rice paddy, bare ground, and built-up classes, the k_{St} roughness coefficients were estimated to be about 13, 22, 18, 25, and $12\,\mathrm{m}^{1/3}\,\mathrm{s}^{-1}$, respectively. As previously mentioned, these values of the k_{St} roughness coefficients were obtained by calibration of the five selected sub-basins.

In brief, the integrated hydrologic and hydrodynamic modeling approach was used to study hydrological behaviors of the selected sub-basins in the upper Nan River basin and to simulate flash flood extents at downstream areas of the sub-basins.

6.3 Spatial index-based approach

In this dissertation a new spatial index-based approach for identifying flash flood hazard areas was developed. The approach is based on the principal component analysis (PCA) and delivers flash flood potential indices (FFPI) from physiographic variables. The determination of these physiographic variables and description of their potentials for flash flood occurrence are given in the following sections.

6.3.1 Derivation of physiographic variables

The computational scheme of the physiographic variables is described in Figure 6.7. To derive the physiographic variables, digital elevation model (DEM), soil, and land use maps were used as input to procedures of terrain pre-processing, stream and catchment delineation. These procedures are explained in Section 6.3.1.1.

6.3.1.1 Terrain pre-processing, stream and catchment delineation

For delineating the stream network and catchment boundary, the DEM was the main input. Normally, a raw DEM contains errors, which break pixel connectivity such as sinks. Thus, the sinks were filled in order to preserve a downward slope along the flow direction. The filled DEM pixels were assigned to values, which indicate flow directions. Subsequently, the outcome of the flow directions was used to compute a flow accumulation in the raster format. The flow accumulation represents how much water is collected from upslope areas and hydrologically contributes to runoff. A value of each pixel in the flow accumulation raster determines the number of upslope pixels that drain into it. Output pixels having high flow accumulation values are the areas of concentrated flow and can be used to define streams. To generate an artificial stream network, a stream delineation threshold was assigned with 1% of the maximum flow accumulation. The flow accumulation pixels, which have values greater than the stream threshold, were defined as a stream. The stream pixels corresponding to the threshold were computed and then linked. Afterwards, the linked pixels were converted to a stream network.

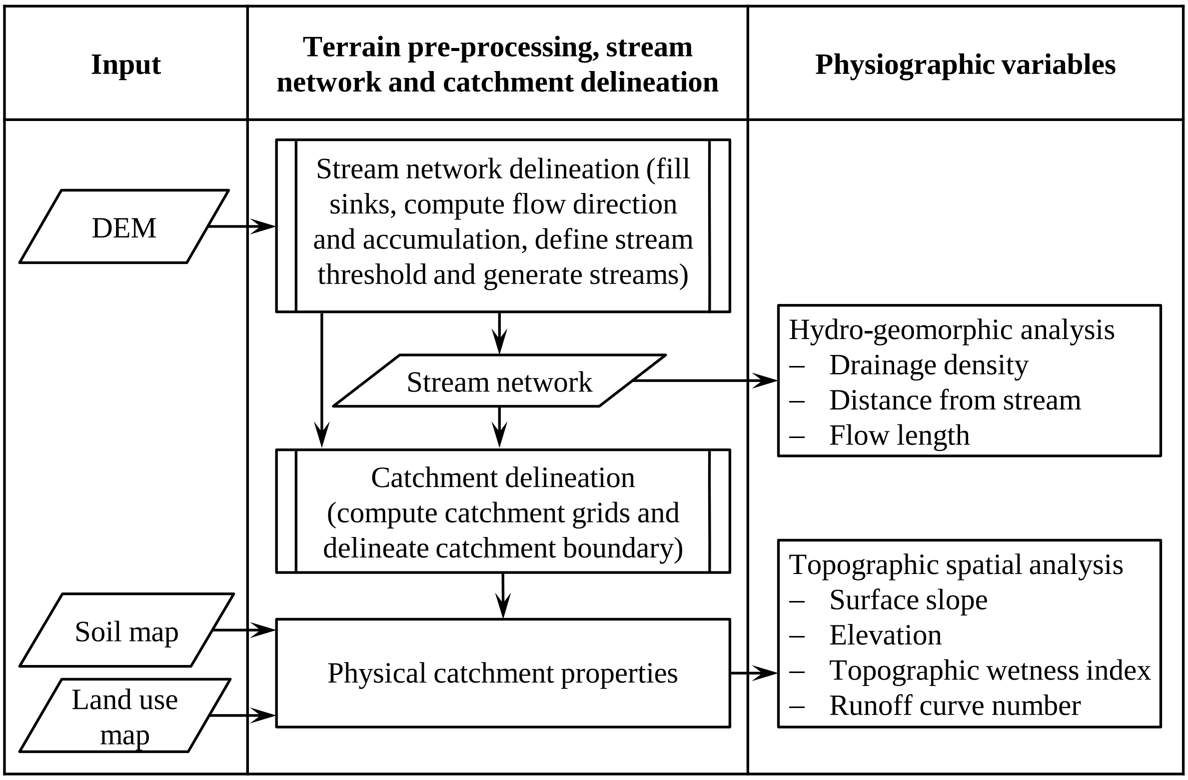


Figure 6.7: Flowchart for deriving physiographic variables: drainage density, distance from stream, flow length, surface slope, elevation, topographic wetness index, and runoff curve number (in ArcMap software).

The catchment delineation was performed using the flow direction pixels and a given catchment outlet. The flow direction pixels, which belong to the outlet, were aggregated to the catchment boundary raster. The catchment boundary feature was delineated from the catchment boundary raster by using the conversion tools of the ArcMap software. The catchment boundary feature was used to clip the soil and land use maps of the Nan River basin.

6.3.1.2 Physiographic variables

The outcomes of terrain pre-processing and stream and catchment delineation were the stream network and physical catchment properties. These characteristics were used to derive the following physiographic variables of the study basin:

- Drainage density (Dd) indicates a draining capacity of stream channels in an area (Horton, 1945). It is defined by the ratio between the total length of stream channels and the area drained by them. Areas with high drainage density tend to have high flood potential because of simultaneous flooding in tributaries (Bell, 1999; Merz and Blöschl, 2005; Pallard et al., 2009) and low infiltration capacity (Patton and Baker, 1976). Conversely, in low drainage density, the flood propagation has longer-delay periods, which allow the main channel to absorb and transmit flows one at a time (Bell, 1999; Kalantari et al., 2014). The drainage density variable is estimated from the stream network by using the Kernel density function (Silverman, 1986).

- Distance from stream (Rbuffer) has been used by many studies (Tehrany et al., 2014; Kazakis et al., 2015; Wang et al., 2015) to identify the flood hazard degree of areas. The areas close to the main stream are likely to have high flood hazard potential. This Rbuffer variable was computed by using the Euclidean distance function in ArcMap.

- Surface slope represents the degree of topographic changes. This variable affects the timing of runoff reaching a given location and the amount of surface runoff and infiltration. A surface with a steeper slope is unlikely to be inundated because of the low capacity to collect water (Wang et al., 2015). In contrast, low-lying lands with gradual slopes are susceptible to flooding because of high infiltration resulting in water logging. In other words, flash floods are generated in upstream areas, where surface slopes are steep and then generally contribute to flooding downstream (Kron, 2002). This surface slope variable was computed from DEM by using the spatial analyst tools in ArcMap.

- Elevation of the ground surface can be directly obtained from DEM. It is the variable that influences runoff generation (Zhang and Montgomery, 1994) and controls the flow direction. Runoff definitely flows from a higher to lower elevation. As a result, lower elevated areas are more likely to be inundated. The lower the elevation, the higher the flood potential in an area is.

- The topographic wetness index (TWI) proposed by Beven and Kirkby (1979) is used to describe the topographic control of hydrologic processes (Sørensen et al., 2006) by indicating the relative properties of saturated conditions at a given site. Saturated areas determined with high TWI values are subject to carry the potential for runoff generation (Woods and Sivapalan, 1997).

- Flow length (FL) measures the distance between a given site and the catchment outlet in the downstream direction. The FL can be used to estimate the time of concentration (El Bastawesy et al., 2009). In a small catchment, short time of concentration is a key factor to generate flash floods as a result of heavy rainfall (Ghoneim et al., 2002; Karamouz et al., 2013). The time of concentration has a high correlation with the FL. Thus, the shorter FL reflects the higher flood potential at the site. The FL variable was calculated from the flow direction data by using the spatial analyst tools in ArcMap.

- Runoff curve number (CN) was introduced by the Soil Conservation Service in 1954 to represent the mean values for land cover, hydrologic soil groups and hydrologic conditions (Rallison and Werner, 1981). The initial CN values can be estimated by integrating soil and land use maps (Chow et al., 1988). High CN values indicate fast runoff formation (Bryndal, 2014). Consequently, areas with high CN values often have a higher flood potential than other areas.

6.3.2 Derivation of flash flood potential index

PCA was applied to establish the FFPI from the aforementioned variables at a spatial resolution of 10 m. Comprehensive description and discussion of PCA can be referred to Dunteman (1989) or Jolliffe (2002). The application of PCA included five main steps:

6.3.2.1 Step 1: normalization of a raw data set

The variables are expressed in different units and have a large variation of value ranges. It would be difficult to evaluate effects of the variables based on their raw values on the FFPI. Consequently, these raw values of each variable were normalized before applying the PCA method in order to exclude the impacts of the different dimensions. The variables were classified into two groups. The first group consisted of drainage density, topographic wetness index, and curve number. These variables increase the degree of flash flood potential (as discussed in the methodology Section 6.3.1.2). In other words, the higher values of these variables indicate the higher flash flood potential. Therefore, the variables in this group were normalized by using Equation (6.1). On the other hand, the variables of the second group decrease the degree of flash flood potential. These variables are slope, elevation, flow length, and distance from stream. They were normalized using Equation (6.2).

$$X_{m,i} = \frac{x_{m,i} - min\left(x_m\right)}{max\left(x_m\right) - min\left(x_m\right)} \tag{6.1}$$

$$X_{m,i} = \frac{max\left(x_m\right) - x_{m,i}}{max\left(x_m\right) - min\left(x_m\right)} \tag{6.2}$$

where $x_{m,i}$ is the raw value, and $X_{m,i}$ is its normalized value of the variable m at the individual input i, $m = 1, 2, \cdots, N_p$, and N_p is the number of the variables, $i = 1, 2, \cdots, N$, and N is the number of the measurements, and $min(x_m)$ is the minimum value, whereas $max(x_m)$ is the maximum value of the variable m.

6.3.2.2 Step 2: calculation of the eigenvalues and eigenvectors of the variable covariance matrix

PCA was applied to transform the normalized values of the variables by using their covariance matrix and then extracted components within the covariance matrix. The transformed values of the variables are called principal component scores (Jolliffe, 2002). Afterwards, patterns of the scores were examined and represented by linear combinations of all variables, which are called eigenvectors or principal components (PCs). Coefficients of the PCs indicate correlations between their PCs and the variables. Each PC explains the variation of the original variables by its variance, which is called an eigenvalue.

6.3.2.3 Step 3: selection of principal components

To benefit a meaningful interpretation of PCA, a set of PCs, which significantly represents the entire data set, needs to be selected. Thus, PCs, which represent low quantities of variance in the data set, were neglected (Dunteman, 1989). Several approaches can be applied to identify how many PCs should be retained in order to effectively summarize the data set (Johnson and Wichern, 2002; Jolliffe, 2002; Rencher, 2002). A common approach proposed by Johnson and Wichern (2002) selected the first *k* PCs, which account for at least 80% or 90% of the variance. In this dissertation, it was found that the first four PCs were responsible for most of the variance. Therefore, *k* was set to four.

6.3.2.4 Step 4: screen of variables

After the number of PCs were selected, it was necessary to screen whether variables were significantly related to the selected PCs. The basin physiographic variables considered here were drainage density, distance from stream (Rbuffer), surface slope, elevation, topographic wetness index (TWI), flow length, and runoff curve number (Figure 6.8). These variables were analyzed and screened by using the PCA method. If a variable has a low correlation with a principal component, it contributes little to an explanation of the variance in the data set. Therefore, a variable, which was low correlated to the PCs, was ignored as it may be redundant with other more predominant variables.

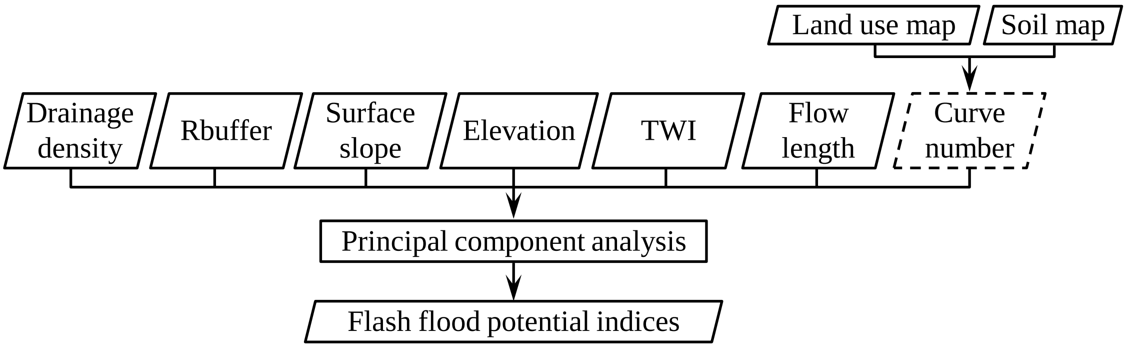


Figure 6.8: Overview of principal component analysis for deriving flash flood potential indices and basin physiographic variables: drainage density, distance from stream (Rbuffer), surface slope, elevation, topographic wetness index (TWI), flow length, and runoff curve number, considered in variable screening procedure.

As previously mentioned in Section 6.3.2.3, the first four PCs were selected for constructing the FFPI. Each PC is a linear combination of variables and its coefficients indicate correlations of the variables with the PC. The curve number variable presented with the dashed line in Figure 6.8 was discarded from the construction of the FFPI according to the PCA method because of its low correlations with the first four PCs. As a result, the remaining variables were maintained and used to compute the FFPI.

6.3.2.5 Step 5: computation of weights for each principal component

The basin physiographic variables were linearly combined and automatically weighted by means of the PCA method in order to derive flash flood potential indices (FFPI). These FFPI were computed by using Equation 6.3. As previously mentioned, it was necessary to select first few PCs, which can be statistically represented most of variances in entire data set. As the first four PCs were selected, the weight (w_i) of each principal component depended on a proportion of the total variance carried by the first four PCs.

$$FFPI = \sum_{i=1}^{4} PCi \cdot w_i \tag{6.3}$$

where

PCi is the i^{th} principal component;
w_i is the weight of each principal component: $w_i = \lambda_i \cdot \left(\sum_{i=1}^{4} \lambda_i\right)^{-1}$;
λ_i is a proportion of the total variance carried by the first four PCs.

6.3.3 Sensitivity analysis of spatial index-based approach

In order to analyze sensitivity of spatial index-based approach, the Nam Hui watershed was selected to demonstrate variations of FFPI according to different spatial rainfall distributions. This watershed was selected because of the flash flood severity in the basin and data availability. On the basis of the local news and oral information, flash floods occurred in the downstream areas of the basin were powerful. They caused riverbank failures, damaged buildings, and swept away animal stalls. The other reason was the rainfall data situation. There were three raingauges located in surrounding the basin as shown in Figure 6.9. This situation allowed to examine the effects of spatial rainfall distributions derived from these raingauges by using different interpolation methods.

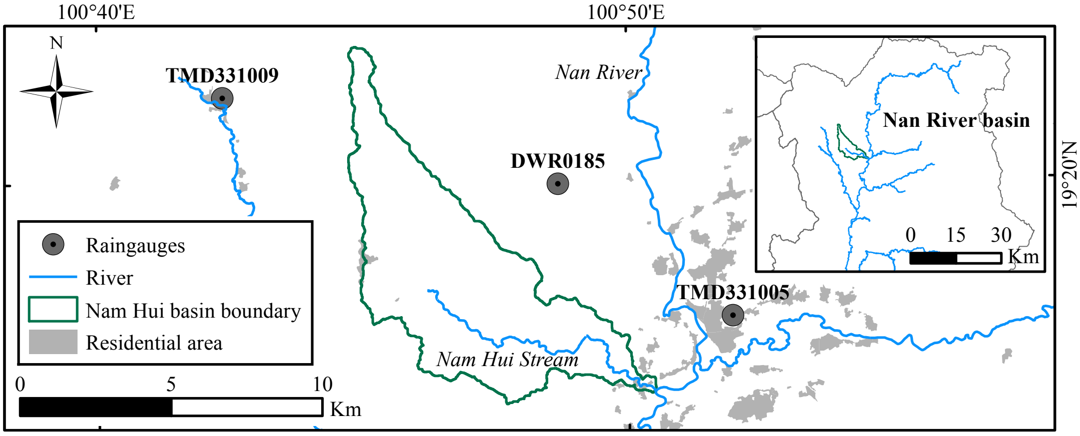


Figure 6.9: The study area (Nam Hui watershed and its location within the upper Nan River basin) and raingauges operated by the Thai Meteorological Department (TMD) and Department of Water Resources (DWR).

In this dissertation, nine experiments were performed for analyzing sensitivity of the spatial index-based approach to different spatial rainfall distributions. These experiments are tabulated in Table 6.2. In the first experiment (Expt. 1), the PCA method linearly combined all the selected physiographic variables – namely, Dd, Rbuffer, surface slope, elevation, TWI, and FL. The other experiments (Expt. 2 - Expt. 9) were conducted for analyzing the sensitivity of the spatial index-based approach to the physiographic variables added by several spatial rainfall distributions.

Table 6.2: Experiments for analyzing sensitivity of the index-based approach to spatial rainfall distributions

Experiment	Spatial rainfall distribution based on	Scenario
Expt. 1	None	-
Expt. 2	radar-rainfall observation	R250611
Expt. 3	inverse distance weighting (3-raingauge)	
Expt. 4	kriging (3-raingauge)	
Expt. 5	Thiessen polygon (3-raingauge)	
Expt. 6	inverse distance weighting (2-raingauge)	
Expt. 7	inverse distance weighting (3-raingauge)	$R_{assumed}$
Expt. 8	kriging (3-raingauge)	
Expt. 9	Thiessen polygon (3-raingauge)	

Note: Physiographic variables, which were drainage density, distance from stream, surface slope, elevation, topographic wetness index, and flow length, were analyzed for all the experiments.

The spatial rainfall distributions were computed according to two rainfall scenarios. The first scenario took rainfall observations during the 25 to 27 June 2011 event (R250611) as the rainfall scenario because of the data availability. As previously mentioned, there were three raingauges located in surrounding the Nam Hui wathershed. However, one of these raingauges (DWR0185) had been operated between 2009 and 2012. As a result, the June 2011 rainfall event was used to set up the first scenario. Different spatial rainfall distributions of this scenario were obtained from radar-rainfall observations and raingauges. The spatial radar-rainfall distribution was used in the Expt. 2. For the Expt. 3, Expt. 4, and Expt. 5, the spatial rainfall distributions of the three day accumulated rainfall were generated from the three raingauges by using inverse distance weighting (IDW), kriging (KG), and Thiessen polygon (TSP) methods, respectively. For the Expt. 6, the spatial rainfall distribution of the three day accumulated rainfall was created from the two TMD raingauges by using the IDW method.

The other rainfall scenario ($\mathrm{R}_{assumed}$) was assumed in order to add variety to the spatial rainfall distributions. Thus, different rainfall depths were assumed to the three raingauges. The higher rainfall depth was given to the raingauge (TMD331009) located close to upstream areas of the Nam Hui watershed. In turn, the smaller rainfall depth was assumed to the raingauge (TMD331005) located close to downstream areas of the watershed. The reason why the raingauge located near to upper parts of the watershed was assigned as a higher rainfall depth was to examine whether the PCA method can manipulate flash flood occurrence because of heavy rainfall. Three day accumulated rainfall depths of 270, 150, and 80 mm, for example, were given to the three raingauges, TMD331009, DWR0185, and TMD331005, respectively. These assumed rainfall depths were spatially interpolated by using the IDW, KG, and TSP methods. The spatial rainfall distribution of the $\mathrm{R}_{assumed}$ scenario generated by the IDW method was used as a climatic variable for the Expt. 7. The spatial rainfall distributions of the $\mathrm{R}_{assumed}$ scenario created by using the KG and TSP methods for Expt. 8 and Expt. 9, respectively. The spatial rainfall distributions for all the experiments are depicted in Section 6.3.3.

6.4 Flash flood hazard mapping

6.4.1 Modeling results to flash flood hazard maps

The integrated modeling approach provided spatial distributions of flood depths and velocities, which were necessary for delineating flash flood hazard maps. For flood hazard rating, Ramsbottom et al. (2006) suggested an empirical formula (Equation 6.4), which was developed in the UK and has been used to estimate flood hazards in many regions (e.g., Kourgialas and Karatzas, 2014; Xia et al., 2011). This formula based on experiments was used to express the degree to which people cannot stand in floodwater because of potential forces caused by flow velocity and deeper water. Therefore, the spatial distributions of flood depths and velocities simulated from the modeling approach were combined by using Equation 6.4 in order to produce flood hazard rating (FHR) grids.

$$\mathrm{FHR} = D(V + 0.5) + \mathrm{DF}, \tag{6.4}$$

where

FHR	is the flood hazard rating [$\mathrm{m}^2\,\mathrm{s}^{-1}$];
D	is the depth of flooding [m];
V	is the flow velocity [$\mathrm{m\,s}^{-1}$];
DF	is the debris factor [$\mathrm{m}^2\,\mathrm{s}^{-1}$].

Ramsbottom et al. (2006) suggested the values of DF depending on the probability that debris would result in a significantly greater hazard for different land uses. Flash floods on urban and built-up areas were the main focus in this dissertation. Therefore, the debris factor (DF) was equal to one when the flood depth was greater than 0.25 m and equal to zero otherwise.

To estimate the degree of flash flood hazard for people in floodwater, flash flood hazards were classified into five classes: very low, low, moderate, significant, and extreme. The very low degree of flash flood hazards was assigned to areas outside flood extent simulated by using the MEADFLOW model. The other classes – namely, low, moderate, significant, and extreme degrees of flash flood hazards, were assigned to the flood extent governed by the MEADFLOW model. To identify flash flood hazard areas, FHR values were classified by thresholds suggested by Ramsbottom et al. (2006). These classifier thresholds for flood hazard rating are summarized in Table 6.3. Low, moderate, significant, and extreme degrees of flash flood hazards were assigned to areas where FHR values were smaller than 0.75, where FHR values were between 0.75 and 1.25, where FHR values were between 1.25 and 2.5, and where FHR values were greater than 2.5, respectively.

Table 6.3: Thresholds for classification of flood hazard rating

Thresholds for flood hazard rating (FHR)*	FHR < 0.75	0.75 < FHR ≤1.25	1.25 < FHR ≤2.5	2.5 < FHR
Hazard degree	Low	Moderate	Significant	Extreme

Note: *Ramsbottom et al. (2006)

6.4.2 Classification of flash flood potential indices

In order to evaluate the spatial index-based approach on flash flood hazard mapping, the FFPI derived by using PCA (described in Section 6.3) needed to be classified and compared with flash flood hazard maps created from the integrated modeling approach. A flowchart of the FFPI classification is shown in Figure 6.10.

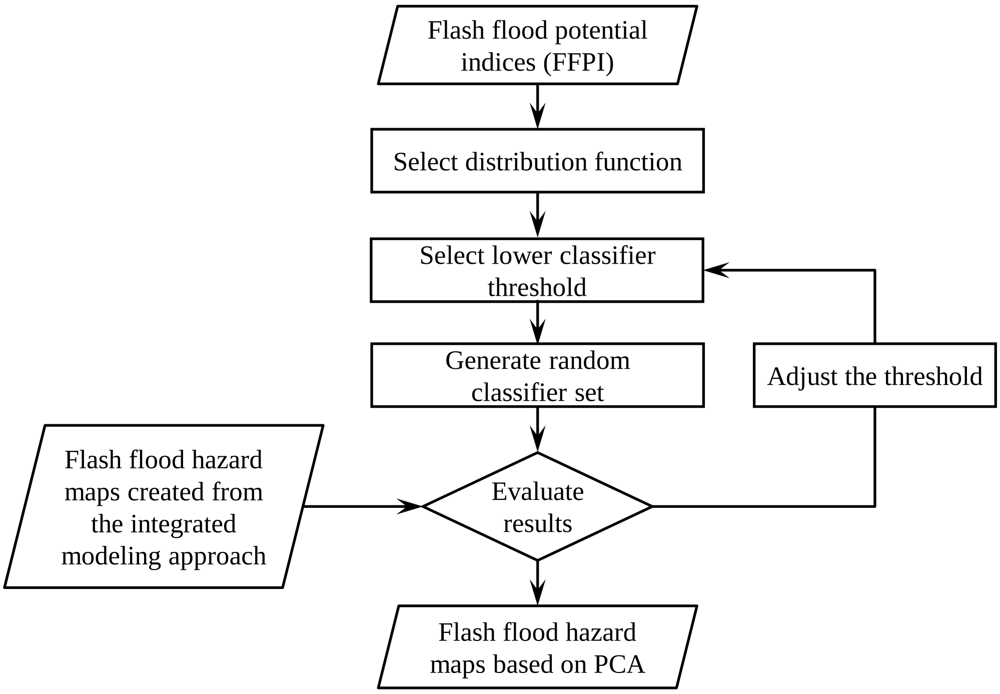


Figure 6.10: Overview of the index classification method based on results of principal component analysis (PCA).

Firstly, the values of the FFPI computed by using PCA were fitted with common distribution functions, such as generalized extreme value, Gumbel and lognormal distributions. The best-fit distribution function of these FFPI values was selected and used to represent the FFPI distribution of the all sub-basins.

Secondly, a lowest classifier threshold was selected in order to indicate areas being considered for the flood hazard classification procedure. The reason of selecting the lowest classifier threshold was to exclude areas with significantly lower flood hazard from the analysis. Thus, the areas indicated by lower values of the FFPI than the threshold were omitted. However, these areas were assigned with very low degree of flash flood hazard. The lowest classifier threshold was selected by trial-and-error. The appropriate threshold would provide a good comparison with the results of the integrated modeling approach. Otherwise, this step would be repeated.

Thirdly, once the lowest classifier threshold was determined, the higher values of the FFPI, which were on the right tail of the selected distribution function, were categorized into four classes. These four classes were used to determine four degree of flash flood hazard: low, moderate, significant, and extreme. To classify the higher values of the FFPI into the four classes, it was necessary to have a set of three classifiers. In this dissertation, 100 000 sets of classifiers were randomly generated. The sets of classifiers were used to classify the FFPI and create flash flood hazard maps.

Finally, flash flood hazard maps created by using the both approaches were compared and evaluated. For performance evaluation, a confusion matrix was applied to compare different classes of flash flood hazard maps created by using the both approaches. On the basis of the confusion matrix, it allowed to statistically measure similarities and differences for all classes. There were three statistical measures – namely, agreement ratio, F-score, and *RMSE*, used to select the most suitable set of classifiers for flash flood hazard mapping. The first two statistical measures were computed with the help of a confusion matrix (described in Section 3.4.2). The other statical measure, *RMSE*, is described in Appendix A.2.1.

6.5 Application of spatial index-based approach in the upper Ping River basin

After a spatial index-based approach and a concept of flash flood potential assessment (FFPA) were developed in the upper Nan River basin, they were applied to four sub-basins of the upper Ping River basin in order to examine and discuss on their suitability for flash flood potential assessment. To apply the spatial index-based approach for identifying areas prone to flash floods, it needed to compute physiographic variables of the sub-basins. These physiographic variables – namely, drainage density, distance from stream, surface slope, elevation, topographic wetness index, and flow length, were computed from a DEM with a 10-m resolution by using the procedures described in Section 6.3.1. Afterwards, the physiographic variables were linearly combined and automatically weighted by the PCA method in order to formulate FFPI.

The concept of FFPA developed in the upper Nan River basin was a classification of FFPI values based on an identified index distribution. This classification required four classifiers to assign five degree of flash flood hazard: very low, low, moderate, significant, and extreme. The very low and low to extreme hazards were separated by a lowest classifier threshold that can be defined by a given percentile of the best-fit probability distribution to the FFPI values of a given watershed. For example, FFPI values of the Nam Mae Hao watershed (Sub-basin P1) located in the upper Ping River basin (Figure 6.11). In the figure, the red dashed line presents a lowest classifier threshold for classifying the FFPI values of the watershed into two zones, very low and low to extreme hazards. For the five selected sub-basins of the upper Nan River basin, very low hazard zones were compared with areas outside flood extent simulated by the integrated modeling approach. On the contrary, low to extreme hazard zones where areas have FFPI values greater than the lowest classifier threshold were compared with the simulated flood extent of the integrated modeling approach. These low to extreme hazard zones were later delineated by a set of three classifier. As a result, low, moderate, significant, and extreme hazard areas were identified and compared with flash flood hazard areas produced by the integrated modeling approach.

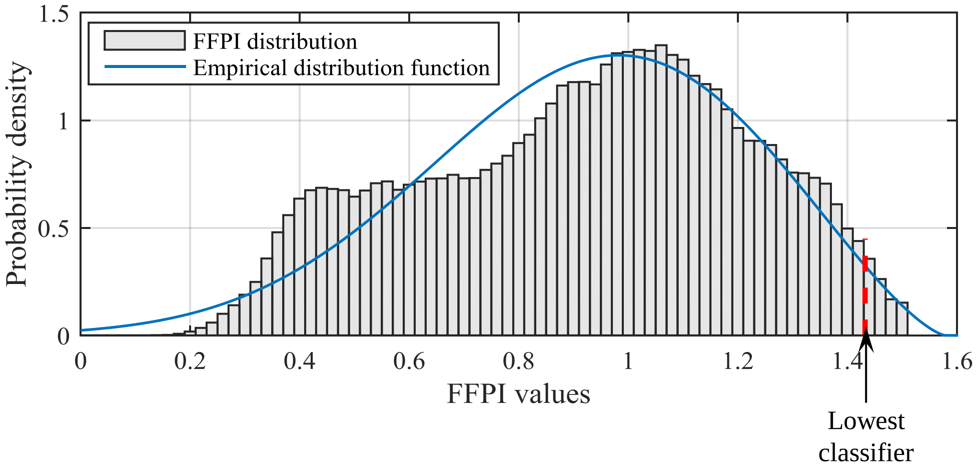


Figure 6.11: Flash flood potential indices (FFPI) distribution based on data of the Nam Mae Hao watershed (Sub-basin P1) in the upper Ping River basin.

As mentioned earlier, the integrated modeling approach for identifying flash flood hazard areas was not applied to the four selected sub-basins of the upper Ping River basin. Therefore, delineation of areas prone to flash floods in these four sub-basins relied on inundated areas observed by the Landsat-7 satellite and historical flood locations gathered from the field investigations. To delineate areas prone to flash floods by using the spatial index-based approach, it was necessary to define a lowest classifier threshold. The lowest classifier threshold that was suitable for the five sub-basins of the upper Nan River was here called the “regional” lowest classifier threshold. In an attempt to examine whether the regional lowest classifier threshold can be directly transferred to use in another regions, it was used to separated areas of the four sub-basins of the upper Ping River basin into very low and low to extreme hazard zones. The latter zone was compared with inundated areas extracted from the Landsat-7 satellite and historical flood locations gathered from the field investigations. The extracted inundated areas were obtained by using a remote sensing technique described in Section 5.5.2.

Flash flood hazard areas in the four sub-basins of the upper Ping River basin were delineated without using any mathematical model, as mentioned above. Thus, flood hazard indicators that are usually formed as combinations of flood depths and velocities were not be obtained to be classified into detailed degree of flood hazards such as low, moderate, significant, and extreme. In order to delineate the low to extreme hazard zones of the four sub-basins, FFPI values of each sub-basins computed by the spatial index-based approach were classified by using common classification methods – namely, natural breaks and geometrical interval classification methods (Dent et al., 2009). These two methods implemented in the ArcMap software grouped the FFPI values into four flood hazard categories: low, moderate, significant, and extreme. Afterwards, their corresponding areas were assigned by four flood hazard rates: low, moderate, significant, and extreme hazards.

6.6 Summary of approach

A major challenge of this research work was dealing with scare and less reliable runoff data to assess flash flood hazard areas. There were two approaches presented here. The first approach was the integrated hydrologic and hydrodynamic modeling approach. Owing to the unavailable runoff data in the study sub-basins, this approach was used to estimate flash flood hazard areas in these sub-basins by modeling the hydrologic response of their parent basin (i.e., the upper Nan River basin). In the upper Nan River basin, there were available hydrological and meteorological data. Therefore, the accuracy of the approach in estimating flash floods of the sub-basins depended on the quality of these data in both spatial and temporal resolutions. High-temporal data such as sub-daily rainfall data were available in the upper Nan River basin. However, the data length was too short for statistical analysis. Daily rainfall data, in turn, were available with long enough periods for the statistical analysis. Therefore, the high-temporal data of radar-rainfall observations, which recorded many extreme rainstorms, were used to derive the temporal rainfall pattern of the upper Nan River basin (i.e., the study region). In addition, the daily rainfall data were used to analyze the probability frequency distribution of the region. This regional probability frequency distribution was combined with the temporal rainfall pattern in order to build synthetic rainfall series. These synthetic rainfall series were input to the integrated hydrologic and hydrodynamic models in order to produce flash flood hazard maps under different rainfall scenarios.

The second approach was a newly developed index-based approach for identifying flash flood hazard. This approach derived flash flood potential indices (FFPI) by means of the principal component analysis (PCA), a multivariate analysis technique. Subsequently, FFPI values were classified into five categories: very low, low, moderate, significant, and extreme degrees of flash flood hazard and flash flood hazard maps then were created. To classify these FFPI values, it was necessary to select classifiers. Suitable classifiers were evaluated with the help of a confusion matrix. This confusion matrix was used to compare areas in different classes of flash flood hazard maps produced by the both approaches.

Both approaches were applied to the upper Nan River basin. The results of these approaches were compared and evaluated by using statistical measures – namely, agreement ratio, F-score, and $RMSE$. The restricted use of the spatial index-based approach could account for areas where detailed flood information is available. This information such as spatial flood depths and velocities can be obtained from postflood ground-based observations or simulated by a mathematical model. However, in areas where the detailed flood information does not exist, the spatial index-based approach needed to be modified.

The spatial index-based approach was applied to the four watersheds of the upper Ping River basin to identify areas prone to flash floods. In these watersheds, there was no detailed flood information. Without this, the comparison of flooded areas was performed differently. Flooded areas identified by the spatial index-based approach were compared with inundated areas extracted from Landsat-7 satellite imagery and historical flood locations gathered from the field investigations.

[This page intentionally left blank]

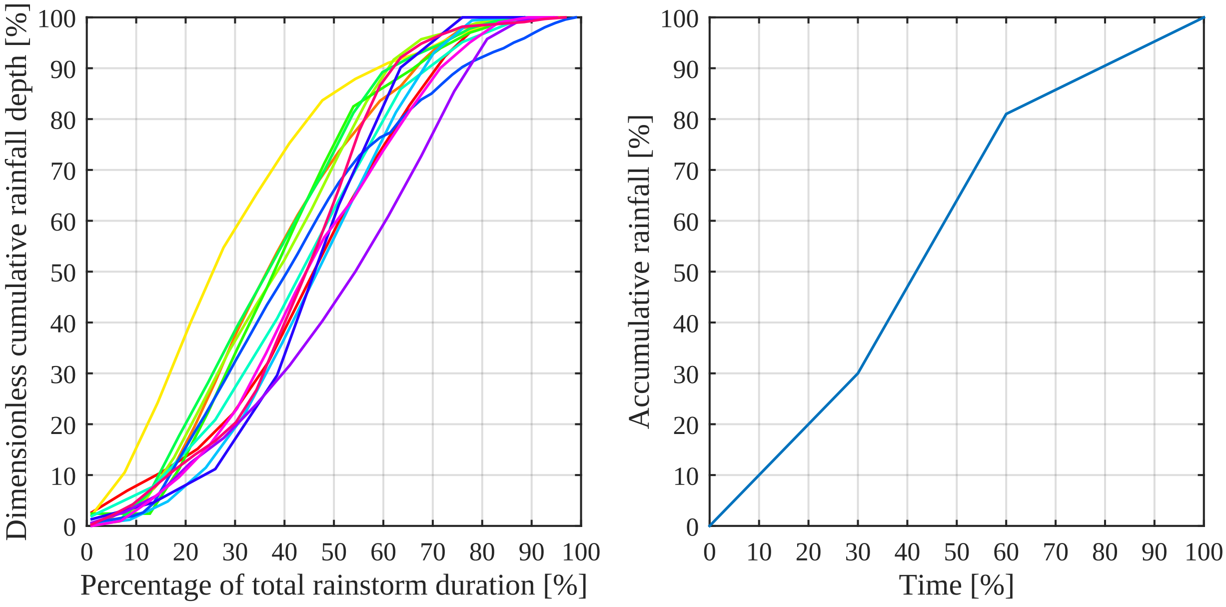
7 Detailed Processing of Work and Results

7.1 Synthetic rainfall series

The development of synthetic rainfall series consists of two main procedures: temporal rainfall pattern and regional frequency analysis.

7.1.1 Analysis of temporal rainfall patterns

The analysis of temporal rainfall patterns was based on radar-rainfall observations at the Chiang Rai weather station. Radar-rainfall images were available every one hour. The temporal patterns of rainfall were characterized by catchment-averaged rainfall series (Yucel, 2015) and based on 13 rainstorms of different rainfall durations in June, July and August of 2008 and 2011. The temporal patterns of rainfall were described in terms of a dimensionless mass curve. The dimensionless mass curve is a plot between dimensionless cumulative rainfall depth and percentage of total rainstorm duration (Figure 7.1). To select the representative temporal pattern of rainfall for this region, the 13 temporal patterns of rainfall were averaged.



(a)

(b)

Figure 7.1: Temporal rainfall patterns of the upper Nan River basin: (a) dimensionless mass curve of 13 rainstorms selected in June, July and August of 2008 and 2011 and (b) representative regional rainfall distribution pattern.

7.1.2 Regional frequency analysis

Regional frequency analysis (RFA) has become necessary when high-resolution rainfall time series are too short for analysis. In this dissertation, sub-daily rainfall data were available for only three or four years but daily rainfall data were available for a long period of 16 years in average. Thus, annual maximum 24 h (AM24H) rainfall data of 11 raingauges were mainly used to carry out the RFA for the upper Nan River basin. Results of the RFA are presented as the following sections.

7.1.2.1 Screening of the data

The AM24H rainfall data at each station was screened by using Grubbs' test (Grubbs, 1969) for outliers. Table 7.1 presents results of the Grubbs's test for two significant levels (α) of 0.05 and 0.01. The test detected an outlier in the AM24H rainfall data of the station 331402. The Grubbs value (G) of the station was greater than the Grubbs' critical values of the both significant levels. Therefore, the maximum value of the AM24H rainfall data in the station 331402 was excluded from the RFA. For the station 331201 it was decided to remain in the RFA because its maximum value was accepted at the 95% confidence level for the Grubbs's test.

Table 7.1: Results of Grubbs' test for annual maximum 24 h rainfall data with a significance level (α) of 0.05 and 0.01

Station	Sample size [year]	G	Grubbs' critical value $\alpha = 0.05$	Grubbs' critical value $\alpha = 0.01$	Max value
090201	18	1.78	2.50	2.82	190.4
331003	19	2.04	2.53	2.85	169.4
331005	19	2.12	2.53	2.85	160.6
331006	19	2.21	2.53	2.85	150.5
331009	18	2.79	2.50	2.82	201.6
331012	7	2.08	1.94	2.10	115.9
331014	7	1.76	1.94	2.10	150.5
331201	20	2.83	2.56	2.88	220.1
331301	19	2.74	2.53	2.85	159.4
331401	19	1.71	2.53	2.85	141.8
331402	15	3.04	2.41	2.71	259.0*

* An outlier

In addition to the individually screening of the rainfall data at each raingauge, it was necessary to examine consistencies of raingauges with the group as a whole. The group consisting of 11 raingauges was tested by the discordance measure (D_i) based on L-moments techniques. Table 7.2 shows the D_i values and L-moment statistics: L-coefficient of variation (L-Cv), L-coefficient of skewness (L-Cs), and L-coefficient of kurtosis (L-Ck) for different raingauges. The D_i values of the 11 raingauges were smaller than 2.63, which is the critical value for 11 numbers of sites given in Hosking and Wallis (1997). This result revealed that no raingauge was discordant. Thus, all raingauges were selected for the RFA.

7.1.2.2 Testing of regional homogeneity

The heterogeneity measures (H_1, H_2, and H_3) for the upper Nan River basin were estimated by performing Monte Carlo simulation experiments from the observed rainfall data of 11 gauges (at-sites). H_1 is the

Table 7.2: Discordance measure and L-moment statistics of AM24H rainfall series in the region

Station	Sample size [year]	Mean	L-Cv	L-Cs	L-Ck	D_i
090201	18	103.4	0.27	0.15	0.00	2.30
331003	19	101.8	0.18	0.21	0.12	0.07
331005	19	97.1	0.18	0.16	0.10	0.22
331006	19	99.7	0.20	0.28	0.05	0.89
331009	18	110.2	0.15	0.36	0.20	0.70
331012	7	71.6	0.16	0.49	0.31	1.55
331014	7	90.1	0.22	0.30	0.02	1.71
331201	20	106.5	0.20	0.27	0.18	0.85
331301	19	88.5	0.16	0.25	0.18	0.35
331401	19	98.4	0.15	0.02	0.00	1.18
331402	14	105.3	0.15	0.01	0.01	1.17

weighted standard deviation of the at-site L-Cvs. H_2 is the average distance between the at-site coordinates and the regional average on an L-Cv versus L-Cs diagram. H_3 is the average distance between the at-site coordinates and the regional average on an L-Cs versus L-Ck diagram, which is called an L-moment ratio diagram. These heterogeneity measures are given in Table 7.3. The heterogeneity statistic of the observed rainfall data from the 11 gauges identified that the basin was a possibly heterogeneous region in an extremely climatic condition because its H_1 value is between one and two, and its H_2 and H_3 values are smaller than one.

Table 7.3: Heterogeneity measures for the upper Nan River basin

Region	Number of stations	H_1	H_2	H_3
Upper Nan River basin	11	1.27**	0.33*	-0.22*

*H_i<1; the region is acceptably homogeneous.
** $1 \leq H_i < 2$; the region is possibly heterogeneous.

7.1.2.3 Identification of the regional distribution

The present study tested the hypothesis of fitting six candidate distributions: Generalized Pareto (GPA), Generalized Extreme Value (GEV), Lognormal (LN), Pearson type III (PE3), Gumbel and Exponential distributions with the series of the AM24H rainfall data by applying the L-moment ratio diagram and the goodness-of-fit measure (Z^{dist}). The L-moment ratio diagram is a plot of L-Cs versus L-Ck (Figure 7.2). The curves illustrate the L-Cs and L-Ck relations for the three-parameter distributions – namely, GPA, GEV, LN, and PE3. For Gumbel and exponential distributions, their L-Cs and L-Ck coordinates present as circle and downward-pointing triangle points, respectively. In addition, the L-Cs and L-Ck coordinates of the at-sites are marked as plus-sign points and their regional mean L-Cs and L-Ck relationship is plotted as a square point. It appears that the regional mean L-Cs and L-Ck relationship is close to the curve of the GPA distribution.

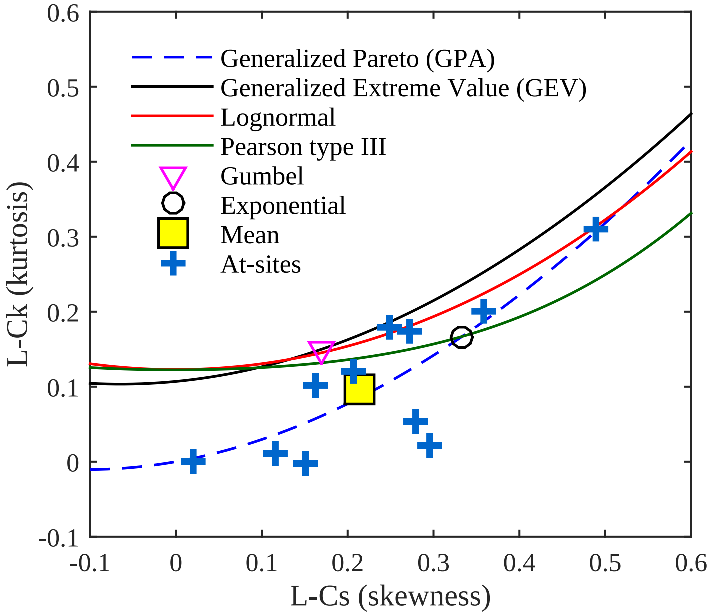


Figure 7.2: L-moment ratio diagram of 11 raingauges (at-sites) and regional rainfall data (mean) for selecting the best-fit frequency distribution for the upper part of the Nan River basin.

Furthermore, the goodness-of-fit measure also confirmed that the GPA distribution was the most suitable probability distribution for the region. According to the goodness-of-fit measure, $|Z^{dist}|$ values of the GPA and PE3 distributions were acceptable at a significance level of 0.10 ($|Z^{dist}| \leq 1.64$) for the AM24H rainfall data of the 11 raingauges as shown in Table 7.4. In contrast, the other distributions failed the hypothesis of the goodness-of-fit measure. As it can be seen in the table, the $|Z^{dist}|$ value of the GPA distribution is smaller than one of the PE3 distribution. Therefore, the GPA distribution was selected to represent the probability distribution of the region.

Table 7.4: The goodness-of-fit measure (Z^{dist}) for the homogeneous region

Distribution	GPA	GEV	LN	PE3
Z^{dist}	-0.47	2.13	1.82	1.20

7.1.2.4 Regional rainfall frequency relationship

On the basis of the L-moment method, the GPA distribution was the most suitable distribution for estimation of rainfall frequency relationships in the upper Nan River basin. The regional rainfall frequency relationships were established by using the GPA parameters and were plotted on probability graphs. Figure 7.3 shows the graphical fitting of the GPA distribution for AM24H rainfall data of the 11 raingauges as well as their regional distribution. In the figure, the scatter points are the reduced variates of the observed AM24H values. The lines represent the relationships between the observed AM24H values and the frequency of occurrence based on the GPA distribution, which was individually estimated by each raingauge. The dashed lines show the relationships between the observed AM24H values and the frequency of occurrence based on the regional GPA distribution, which was estimated by using regional GPA parameters of the 11 raingauges.

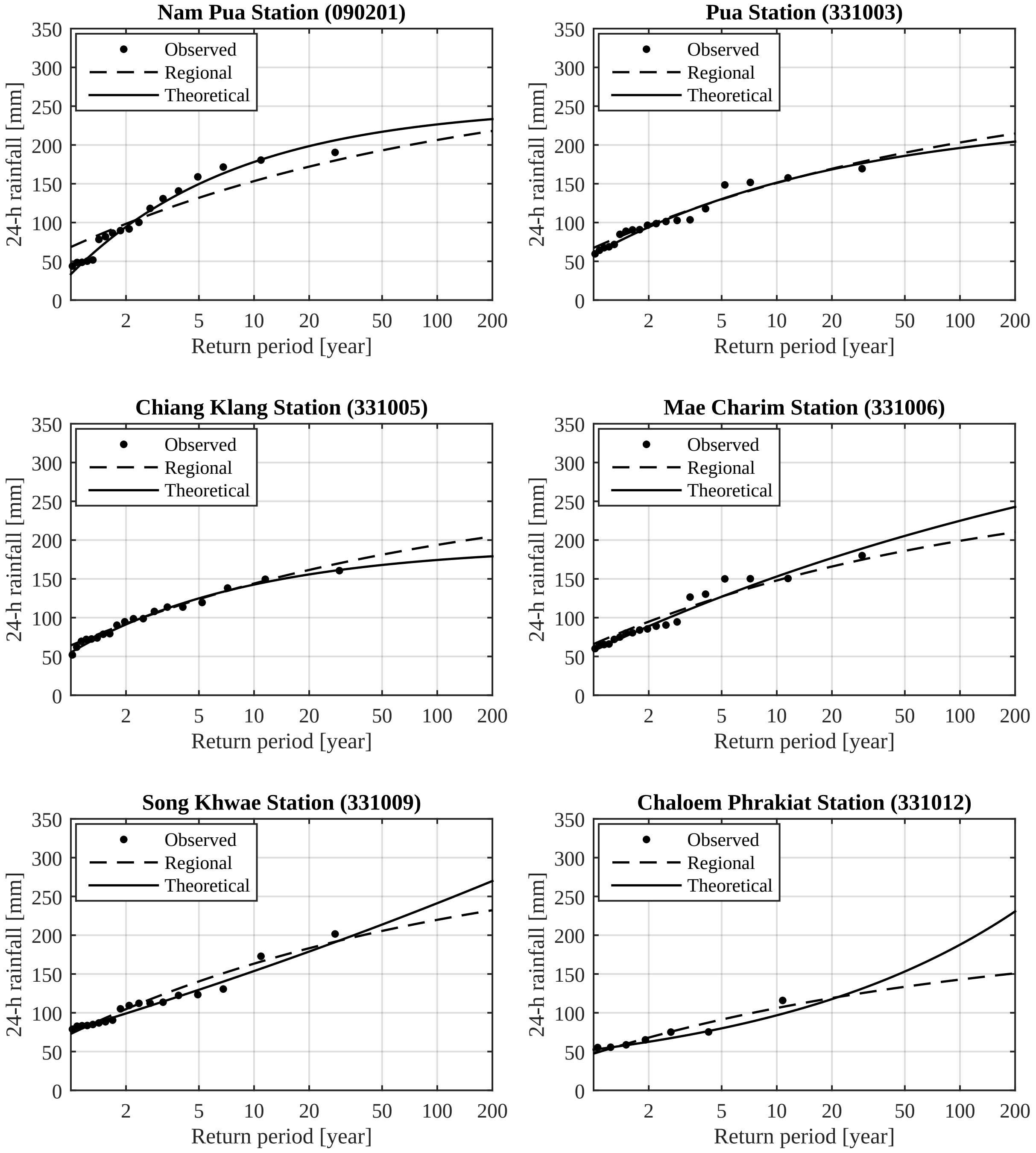


Figure 7.3: Regional and generalized Pareto (GPA) frequency curves fitted to annual maximum 24 h (AM24H) rainfall data of the upper Nan River basin: (points) reduced variates of observed AM24H values, (dashed line) regional GPA frequency curve estimated by regional GPA parameters of 11 raingauges, and (line) GPA frequency curve individually estimated by each raingauge.

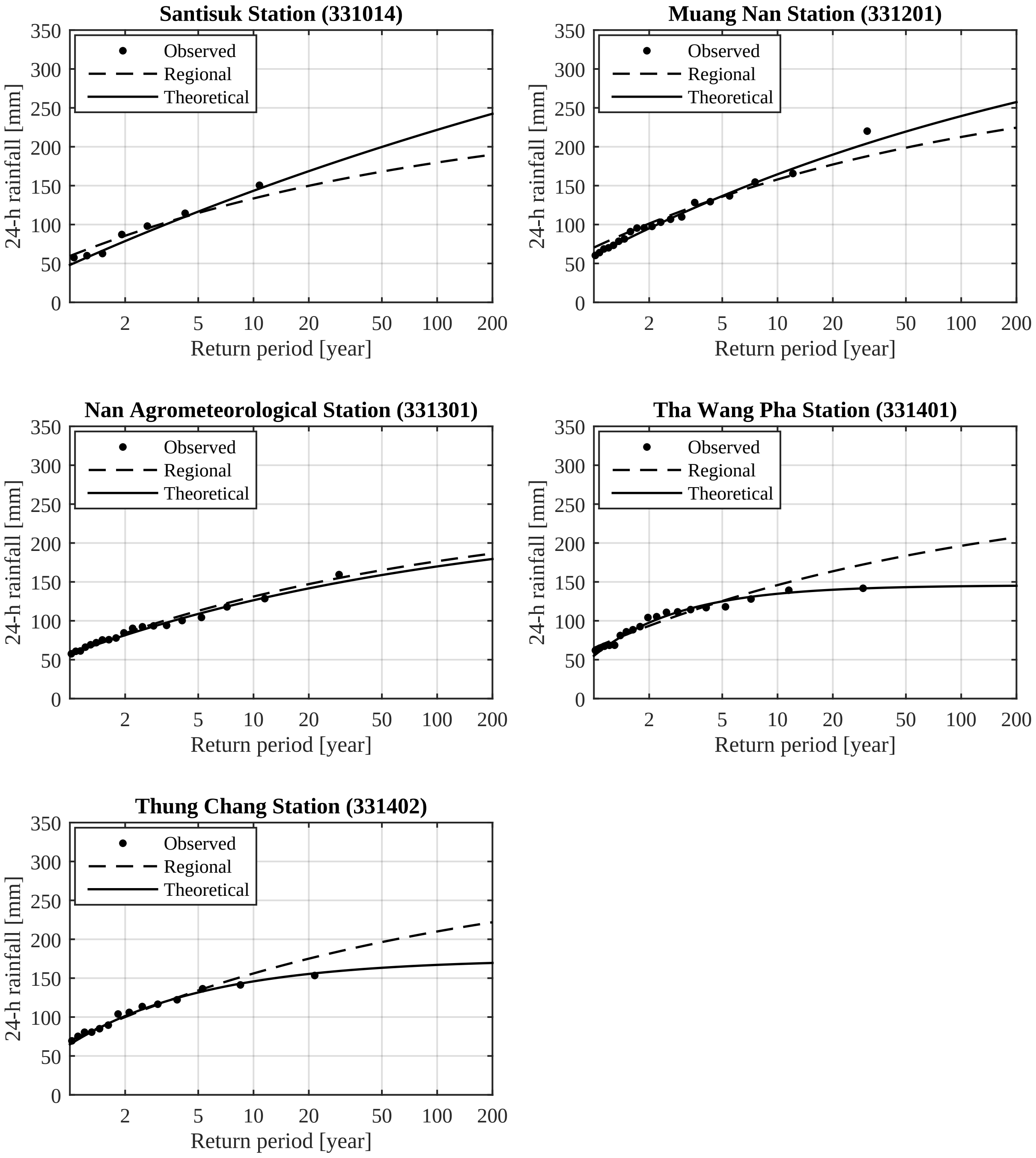


Figure 7.3: Regional and generalized Pareto (GPA) frequency curves fitted to annual maximum 24 h (AM24H) rainfall data of the upper Nan River basin: (points) reduced variates of observed AM24H values, (dashed line) regional GPA frequency curve estimated by regional GPA parameters of 11 raingauges, and (line) GPA frequency curve individually estimated by each raingauge (continued).

For design purposes, the regional rainfall frequency relationships were combined with Bell/SCS type II ratios (Bell, 1969; Awadallah et al., 2011) in order to create depth-duration-frequency (DDF) curves (Figure A.1 in Appendix A.3.1) as described in Section 6.1. Afterwards, these DDF curves were integrated with the regional temporal rainfall pattern from the section above (Section 7.1.1) to form synthetic rainfall series (e.g., Figure 7.6) that were used for estimating design flash floods in hydrologic modeling (see Section 7.2.1).

7.2 Integration of hydrologic and hydrodynamic modeling

7.2.1 Application of PANTA RHEI hydrologic model

In order to characterize flash floods in the ungauged sub-basins of the upper Nan River basin, it was necessary to analyze historical extreme flood events, which were monitored in their parent basin (upper Nan River basin). In the last decade, downstream areas of the parent basin damaged by two extreme flood events in August 2008 and June 2011. Thus, these events were selected for constructing the PANTA RHEI hydrologic model.

The PANTA RHEI model acceptably simulated the two extreme flood events. The 2011 flood event between 25 and 29 June 2011 was used to calibrate the PANTA RHEI model. For validating the model, the period between 17 July and 11 August 2008 was selected. Figure 7.3 shows the results of the PANTA RHEI model using the disaggregated hourly rainfall data during the periods of calibration and validation. To evaluate the results, the present study used two performance criteria, which were efficiency index (*EI*) and relative peak error (*RPE*). The results of PANTA RHEI, which were simulated hydrographs, were compared with hydrographs observed at the river gauging stations: N.1, N.64, and N.65. On the basis of performance criteria, the simulated and observed hydrographs illustrated in Figure 7.3 show good simulation for the two extreme flood events considered. For example, the *EI* values for calibration and validation at the river gauging station N.64 were 0.95 and 0.81, respectively, which means that the PANTA RHEI model was capable to reasonably reproduce extreme flood events. Moreover, the *RPE* measure indicated a slight deviation between simulated and observed flood peaks by -0.24%, which means that the simulated flood peak was slightly lower than the observed one, in the calibration. However, the PANTA RHEI model slightly overestimated flood peak by 7.0% in the validation. According to the visual comparison with the observed hydrographs and the performance criteria, the PANTA RHEI model was able to reproduce flood peak and volume. Thus, this PANTA RHEI model was considered satisfactory and suitable for simulating floods in the parent basin.

7.2.1.1 Sensitivity analysis of model parameter sets

The purpose of sensitivity analysis was to examine uncertainty because of model parameters. Figure 7.4 shows bandwidth of simulated hydrographs based on the examined parameters in the five sub-basins: the Nam Rim watershed (N1), the Nam Hui watershed (N2), the Nam Lae watershed (N3), the Nam Pua watershed (N4), and the Nam Khwang watershed (N5). As a result of runoff simulation within the boundaries ($\pm 20\%$) of the calibrated parameters, the upper and lower bandwidth boundaries of simulated hydrographs were higher and lower, respectively, than the simulated hydrographs (red solid lines) based on the calibrated parameters (reference hydrographs) as shown in Figure 7.4. According to this sensitivity analysis, relative differences of peaks between the upper or lower bandwidth boundaries of simulated hydrographs and the reference hydrograph were about $\pm 7\%$. This indicates that simulated runoff peaks were insensitive to variations in model parameters. Thus, the calibrated parameters of the station N.64 could be transferred to simulate flash flood hydrographs of the five sub-basins.

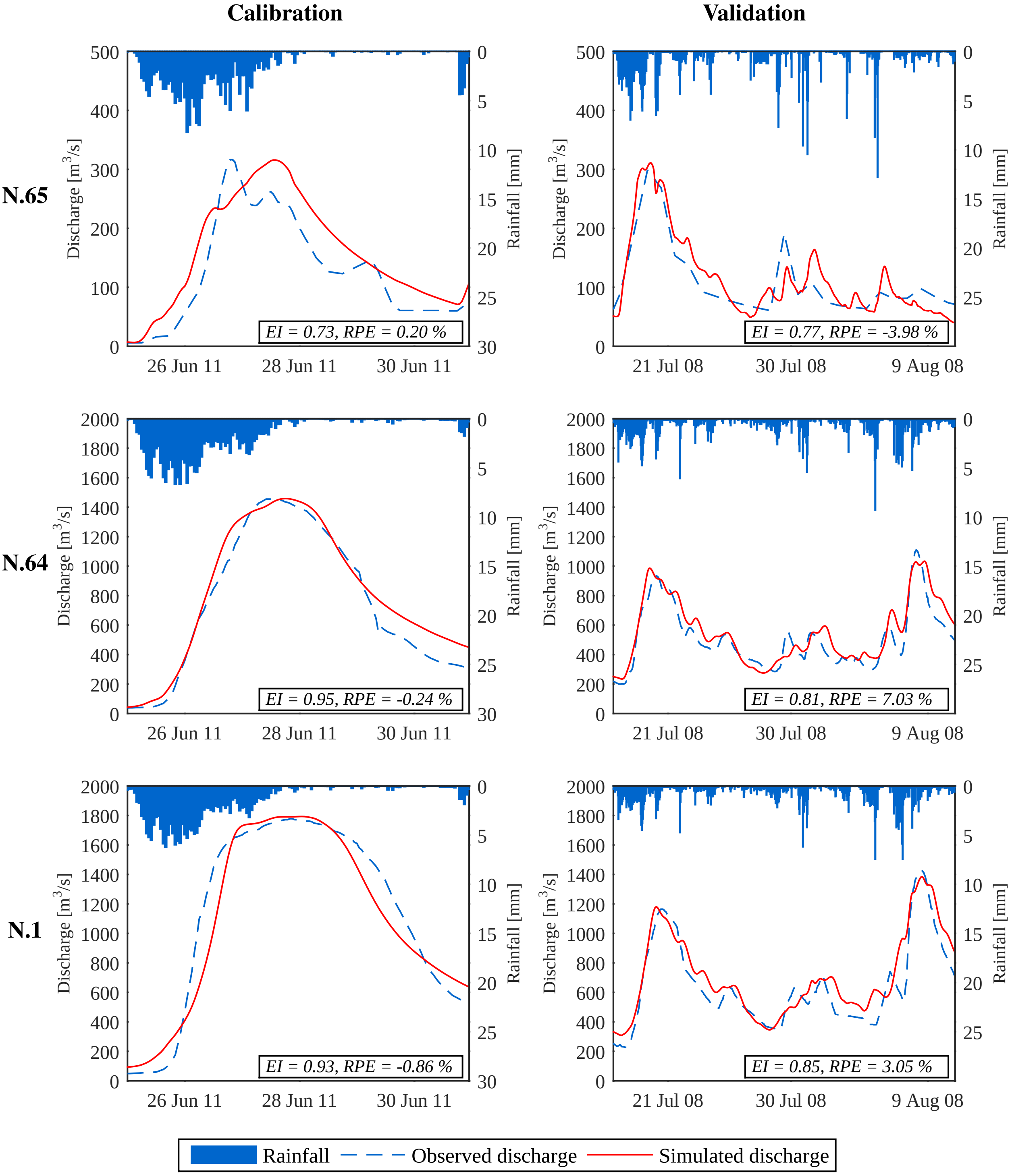


Figure 7.3: PANTA RHEI simulation results (red line) compared to recorded discharge (dash blue line) at Ban Pang Sa (N.65), Ban Pha Khwang (N.64) and the Forestry Office, Nan (N.1): (Left) calibration for the 25-30 June 2011 event, (Right) validation for the 17 July-11 August 2008 event.

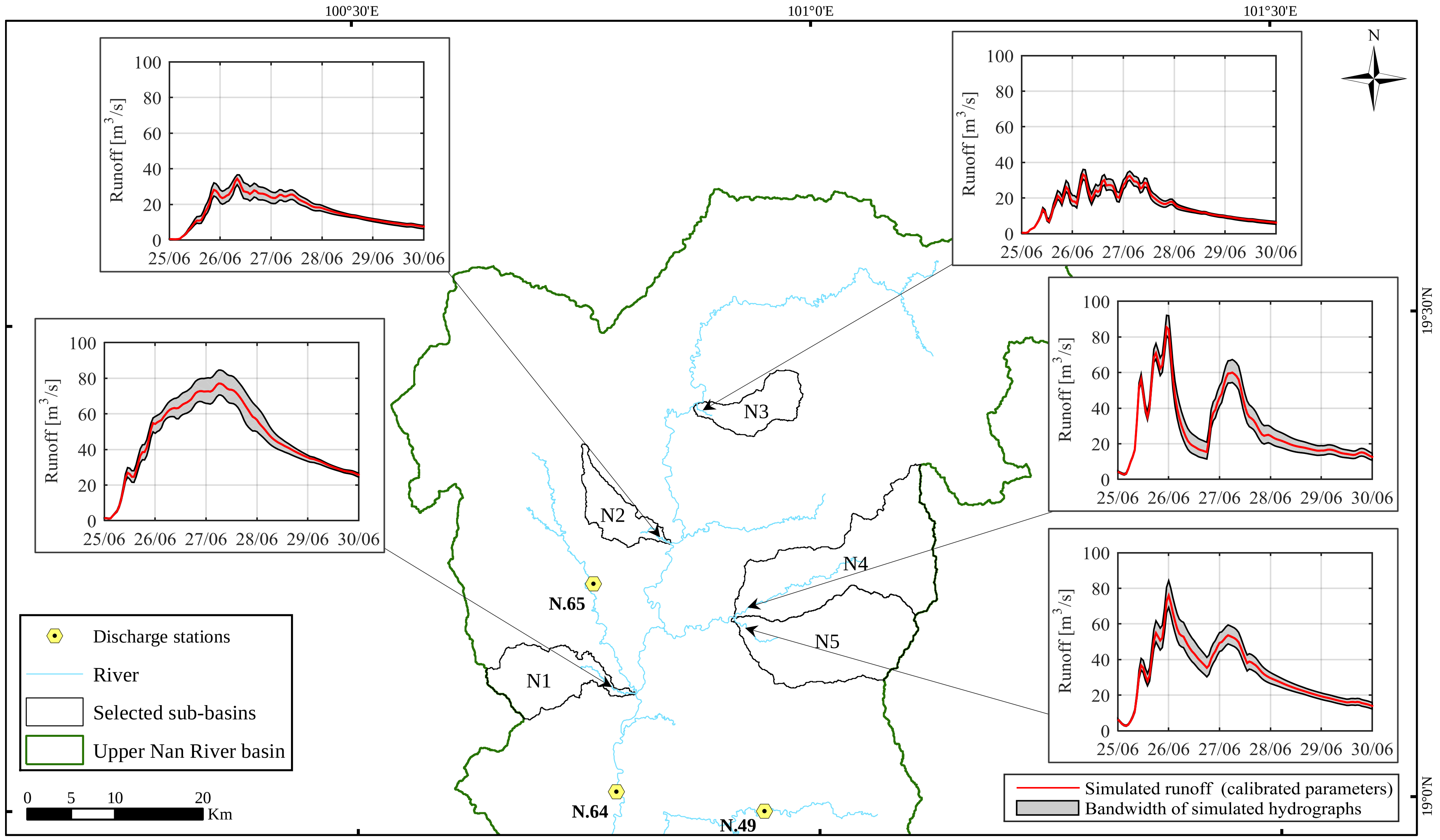


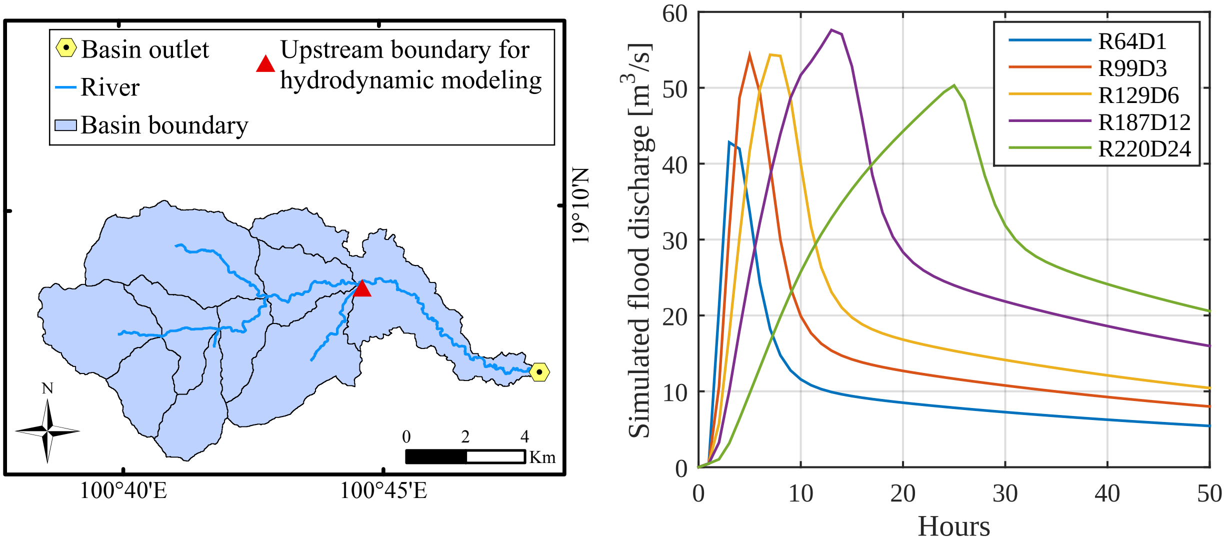
Figure 7.4: Sensitivity analysis results of the combined parameter sets within the boundaries between -20% and +20% of the calibrated parameters for the five sub-basins compared with simulated floods in 2011 (Nam Rim watershed (N1), Nam Hui watershed (N2), Nam Lae watershed (N3), Nam Pua watershed (N4), and Nam Khwang watershed (N5)).

7.2.1.2 Analysis of historic extreme rainfall events

The occurrence of flash floods in small to midsize watersheds depends on a combination of rainfall amount and duration. Therefore, it was necessary to analyze extreme rainfall events observed in the study region. Table 7.5 shows observed extreme rainfall events of the upper Nan River basin. These rainfall events were assumed to be uniformly distributed over the sub-basin and were transformed to flash flood hydrographs at interior locations of the sub-basin by using the PANTA RHEI model. For example, an interior location of the Nam Rim watershed (N1), which was a upstream boundary for hydrodynamic modeling, is shown in Figure 7.5a and simulated flash flood hydrographs with corresponding to R64D1, R99D3, R129D6, R187D12, R220D24 events are shown in Figure 7.5b. For example, the R64D1 event represents a rainfall scenario that is the maximum rainfall amount of approximately 64 mm falling in a duration of one hour.

Table 7.5: Historic extreme rainfall events observed in the upper Nan River basin

Rainfall event	Maximum point rainfall amount [mm]	Rainfall duration [hour]
R64D1	64.4	1
R99D3	98.9	3
R129D6	129.1	6
R187D12	187.0	12
R220D24	220.1	24



(a) (b)

Figure 7.5: Simulated flash flood hydrographs for the Nam Rim watershed (N1): (a) basin boundary and location of flash flood hydrographs (b) simulated flash flood hydrographs for extreme rainfall events.

As it can be seen in Figure 7.5b, the Nam Rim watershed had quick responses to extreme rainfall. The simulated peak discharge with corresponding to the maximum rainfall amount of 187 mm falling in a duration of 12 h (R187D12) was the highest peak among ones simulated with corresponding to the five rainfall events. The second highest peak discharge was computed from the rainfall scenario of the maximum rainfall amount about 129 mm falling in a duration of six hours (R129D6). From this analysis of extreme rainfall events, the rainfall amounts between 130 and 200 mm falling in durations of 6 to 12 h along with soil moisture conditions have the potential to cause a flash flood in the study region.

7.2.1.3 Estimation of design flash floods

The design flash flood events at interior locations of the selected sub-basins were estimated depending on the calibrated model parameters of the parent basin. For the estimation of the design flash flood events, it was necessary to determine a critical storm duration. Depending on observed rainfall data in the study region, a storm duration being longer than the concentration time of the sub-basin can cause flash floods. This agrees with the explanation of Ponce (1989) that runoff occurs mainly as overland flow rather than as streamflow when storm duration exceeds concentration time of a given sub-basin. Thus, the critical storm duration of six hours was chosen for the design floods because the concentration time of the selected sub-basins was shorter than six hours.

As an example, synthetic six-hour rainfall series for 10-, 20-, and 50-year return periods (RTr10, RTr20, and RTr50, respectively) at the Tha Wang Pha meteorological station (331401) are plotted, in Figure 7.6. The synthetic six-hour rainfall series for different return periods were feed into the PANTA RHEI model and their corresponding flood hydrographs were estimated.

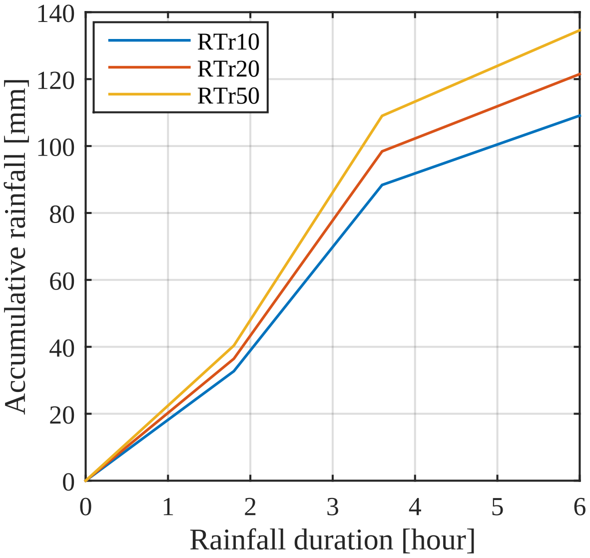
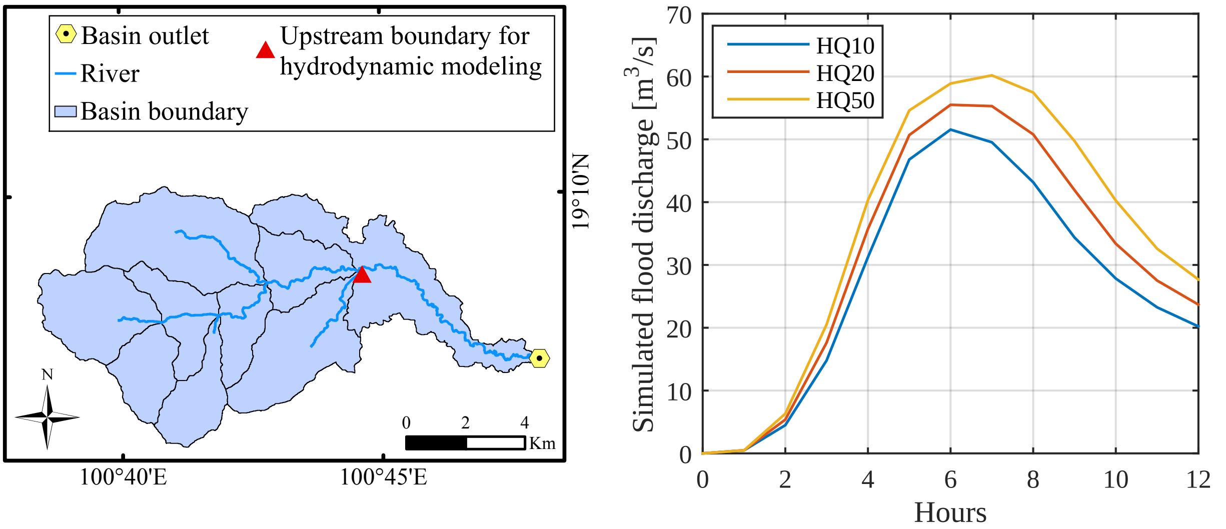


Figure 7.6: Synthetic rainfall series based on regional frequency analysis at the Tha Wang Pha meteorological station (331401).

As previously mentioned, synthetic rainfall series of all the 11 daily raingauges were spatially distributed by the IDW method within the PANTA RHEI model. The purpose of using the temporal and spatial variations of rainfall based on the regional frequency analysis of extreme rainfall was to create synthetic regional rainfall scenarios for the sub-basins of the upper Nan River basin. These regional rainfall scenarios such as rainfall scenarios of RTr10, RTr20, and RTr50 were used as inputs to the PANTA RHEI model in order to simulate design flood hydrographs at interior locations of a given watershed. For example. Figure 7.7a depicts an interior location of the Nam Rim watershed (N1), which was an upstream boundary for hydrodynamic modeling, and Figure 7.7b shows design flood hydrographs (HQ10, HQ20, and HQ50) for the regional rainfall scenarios of RTr10, RTr20, and RTr50, respectively.



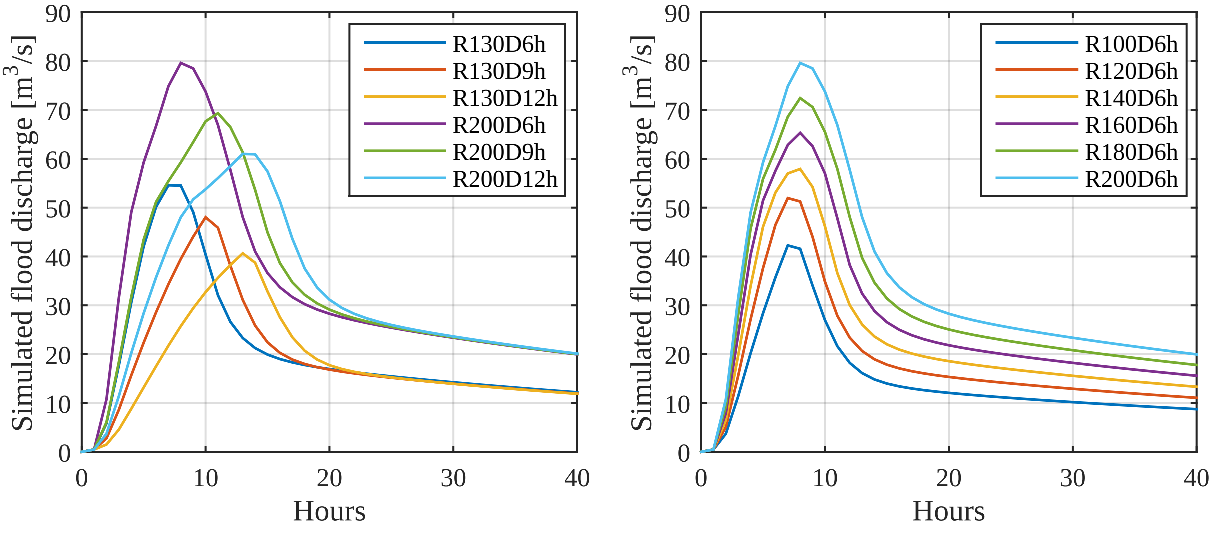
(a) (b)

Figure 7.7: Design flash flood hydrographs for the Nam Rim watershed (N1): (a) basin boundary and location of flash flood hydrographs (b) design flash flood hydrographs responding to rainfall return periods (10-, 20-, and 50-years return period).

On the basis of historical rainfall observations and analysis of extreme rainfall events in the Section 7.2.1.3, it can be noticed that the rainfall amounts between 130 and 200 mm falling in durations of 6 to 12 h along with soil moisture conditions have the potential to cause a flash flood in small to midsize watersheds of the study region.

Therefore, several combinations of rainfall amounts and durations served as an input into the PANTA RHEI in order to estimate hydrographs for design flash floods. For example, rainfall amounts of 130 and 200 mm were assumed to be uniformly distributed over the Nam Rim watershed for durations of 6, 9, and 12 h. Their corresponding flash flood hydrographs at the interior location of the watershed are shown in Figure 7.8a.

According to Figure 7.8a, a rainstorm with the same rainfall amount falling in shorter duration has a higher flood peak. Therefore, the duration of six hours was considered to be the critical storm duration of the watershed and was used for design flash flood hydrographs. For instance, Figure 7.8b shows flash flood hydrographs at the interior location of the Nam Rim watershed.



(a)

(b)

Figure 7.8: Design flash flood hydrographs for the Nam Rim watershed (N1): (a) hydrographs with corresponding to rainfall events with amount of 130 and 200 mm (b) hydrographs with corresponding to rainfall events with amount of 100, 120, 140, 160, 180 and 200 mm within a six-hour period.

The reason why flash flood hydrographs were simulated at the interior location of the watershed was that their peaks were used to assign as a upstream boundary for hydrodynamic modeling. As a result, the flood peaks based on synthetic rainfall series and uniform rainfall distributions were routed by using the MEADFLOW hydrodynamic model in order to create flash flood hazard maps under different rainfall scenarios.

7.2.2 Hydrodynamic modeling

The MEADFLOW hydrodynamic model was applied to the downstream areas of the five selected sub-basins to estimate flash flood extents and depths. The estimated flood depths of MEADFLOW were compared with high-flood marks, which occurred during the 2011 flood event as shown in Figure 7.9. The information of the high-flood marks were collected during the fieldwork of this dissertation by interviewing flood witnesses and reviewing flash flood report. As it can be seen in Figure 7.9, the estimated flood depths were higher than the observed flood marks in some locations. Differences between the estimations and observations may be caused by errors in post-flood measurement, errors in flood peak simulation of the hydrologic model, errors due to simplifications of the model structure, and changes in land elevation. However, the overall results were acceptable according to two statistical measures – namely, *EI* and *RMSE*. The *EI* value was 0.69 and the *RMSE* was 0.17 m for the comparison between observed and estimated flood marks at downstream areas of the five selected sub-basins in the upper Nan River basin.

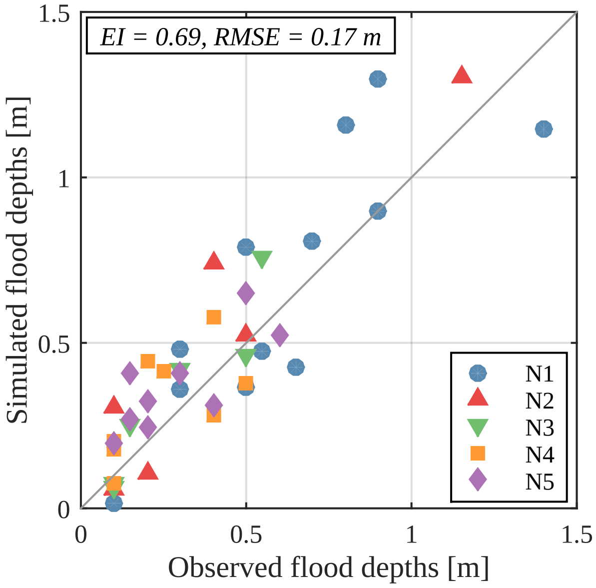
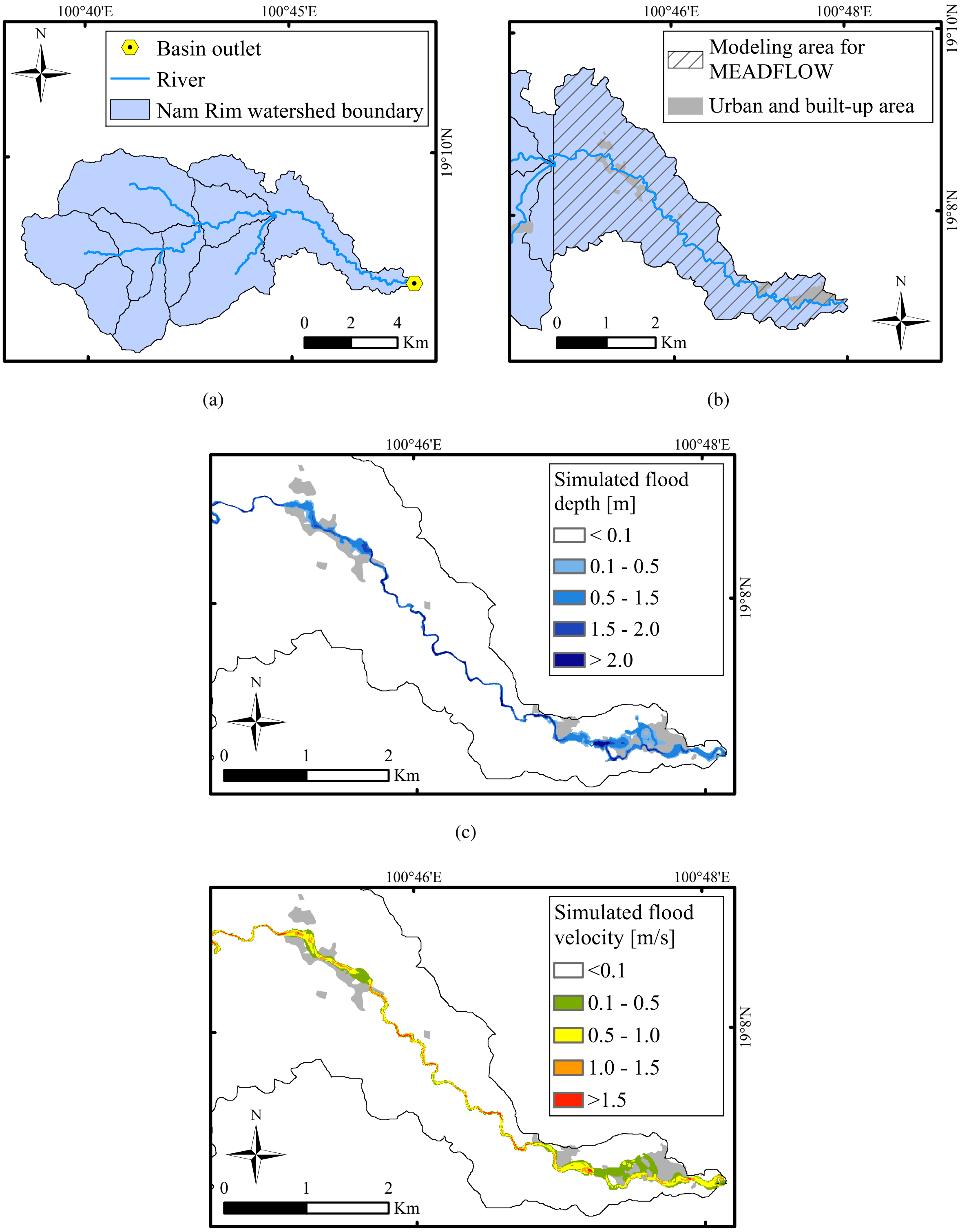


Figure 7.9: Comparison between maximum flood marks, which occurred in 2011, and simulated flood depths calculated by using the MEADFLOW model in the Nam Rim watershed (N1), the Nam Hui watershed (N2), the Nam Lae watershed (N3), the Nam Pua watershed (N4), and the Nam Khwang watershed (N5).

MEADFLOW was applied to estimate flash flood extents, depths, and velocities at downstream areas of the study sub-basins. For example, Figure 7.10a shows the Nam Rim watershed (N1) boundary and Figure 7.10b shows the downstream areas of this watershed, which were used for hydrodynamic modeling. The upstream boundary of the MEADFLOW model was obtained from simulated flood peaks of the PANTA RHEI model. Figure 7.10c and 7.10d show flood depths and velocities under a regional rainfall scenario of the10-year return period, respectively.

Afterwards, these flood depths and velocities were combined in order to delineate flash flood hazard areas (as presented in Section 7.4.1). These flash flood hazard areas were later used for comparison with results obtained by the spatial index-based approach in Section 7.3.



(a)

(b)

(c)

(d)

Figure 7.10: Nam Rim watershed: (a) Nam Rim watershed boundary, (b) modeling area for MEADFLOW, (c) simulated flood depths under a regional rainfall scenario of RTr10, and (d) simulated flood velocities under a regional rainfall scenario of RTr10.

7.3 Spatial index-based approach

With reference to Section 6.3, flash flood potential indices (FFPI) were derived by using the principal component analysis (PCA) of physiographic variables. These variables were drainage density (Dd), distance from stream (Rbuffer), surface slope, elevation, topographic wetness index (TWI), flow length (FL), and runoff curve number (CN). On the basis of PCA, the variables were linearly combined and automatically weighted to form the FFPI, which can be used to identify areas prone to flash floods.

The spatial index-based approach was developed in the five selected sub-basins of the upper Nan River basin. For demonstrating results of this spatial index-based approach, the Nam Hui watershed (N2) was selected because of the flash flood severity in the basin and data availability. According to the local news and oral information, flash floods occurred in the downstream areas of the basin were severe. They caused riverbank failures, damaged buildings, and swept away animal stalls. The other reason was the rainfall data situation. There were three raingauges located in surrounding the basin. This situation allowed to examine the effects of spatial rainfall distributions derived by using different interpolation methods.

7.3.1 Variable screening

In the initial PCA analysis, there were seven physiographic variables. These variables were computed from the DEM of the Nam Hui watershed (N2). Table 7.6 shows the results of PCA, which are eigenvectors and variance of all principal components (PCs) for the watershed. The eigenvectors, which are PCs, are represented in columns. In the table, the values of variables (rows), which are called component coefficients, illustrate the correlations between the variables and their PCs. High values of the component coefficients were highlighted in bold to emphasize the high relative importance of the variables to their PCs. Moreover, the variance of the PCs indicates the amount of variation, which is accounted for by each PC, in the whole data set.

Table 7.6: All principal components for screening variables of the Nam Hui watershed (N2) and their variance

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.11	**0.98**	0.03	-0.05	0.06	-0.04	0.11
Rbuffer	**0.55**	-0.08	-0.26	-0.12	**0.52**	**-0.52**	-0.27
Surface slope	0.13	-0.07	**0.95**	-0.17	0.06	-0.22	0.00
Elevation	**0.55**	0.05	0.06	0.12	**-0.51**	0.31	**-0.57**
TWI	0.04	0.02	0.14	**0.96**	0.21	-0.07	0.04
FL	**0.61**	-0.14	-0.04	-0.02	-0.08	0.22	**0.75**
CN	0.02	0.00	0.11	-0.10	**0.65**	**0.73**	-0.16
Variance [%]	52.30	19.88	13.39	5.93	5.18	2.77	0.54
Cumulative variance [%]	52.30	72.18	85.57	91.50	96.68	99.46	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

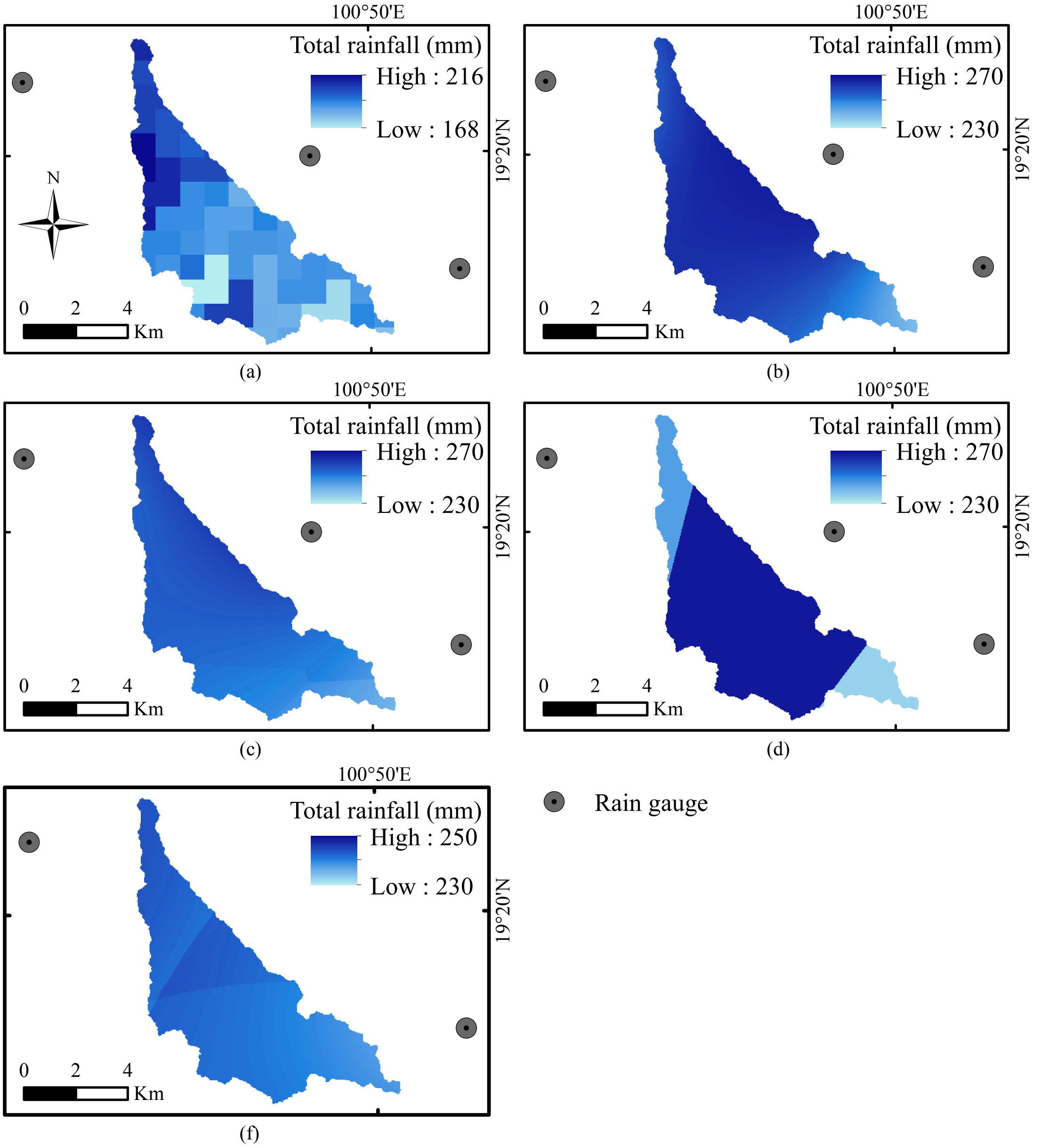
To achieve a general purpose of PCA, a first few PCs, which capture the majority of the variance in the data set, were selected. In this variable screening procedure, it was found that the first four PCs represented 91.5% of the total variance.

In addition to the selection of the first four PCs, PCA was applied to screen the seven variables by exploring the interrelations among them and considering their correlations with the selected first PCs. As it can be seen in Table 7.6, the coefficients of PC1 are positive for all the variables. It means that all the variables are positively correlated among themselves. In this data set, PC1 had large positive associations with Rbuffer, elevation, and FL. However, both positive and negative signs were found in the coefficients of the PC2. For example, the coefficients of PC2 associated with Dd, elevation, TWI, and CN were positive and the other coefficients were negative. While the positive sign indicates positive correlations with a given principal component, the negative sign simply means the reverse. According to the PCA theorem, the coefficients of the second and remaining PCs are expected to be positive and negative when the coefficients of PC1 are all positive. In other words, the first principal component must be uncorrelated and orthogonal to the other principal components. In the selected first four PCs, their coefficients, which indicated their correlations with the CN variable, were very low. Moreover, it was found that this CN variable had low correlations with the first four PCs not only in the Nam Hui watershed (N2), but also in the other selected watersheds of the upper Nan River basin. Therefore, the CN variable was removed from the spatial index-based approach.

In short, it was found that significant physiographic variables for deriving the FFPI were Dd, Rbuffer, surface slope, elevation, TWI, and FL.

7.3.2 Sensitivity analysis to spatial rainfall distributions

For sensitivity analysis of the spatial index-based approach, nine experiments (Expt. 1–Expt. 9) were performed as described in Section 6.3.3 and tabulated in Table 6.2. The first experiment (Expt. 1) used the PCA method to derive a FFPI distribution from the selected physiographic variables without considering rainfall distributions. The other experiments (Expt. 2–Expt. 9) included physiographic variables with different spatial rainfall distributions obtained from radar-rainfall observations and raingauges into the PCA method. The spatial radar-rainfall distribution was used in the Expt. 2. For the Expt. 3, Expt. 4, and Expt. 5, the spatial rainfall distributions under the R250611 scenario were generated from three raingauges located in surrounding the basin by using inverse distance weighting (IDW), kriging (KG), and Thiessen polygon (TSP) methods, respectively. For the Expt. 6, the spatial rainfall distribution under this rainfall scenario was created from the two TMD raingauges by using the IDW method. Different spatial rainfall distributions of the rainfall scenario obtained from radar-rainfall observations and raingauges are depicted in Figure 7.11. For the Expt. 7, Expt. 8, and Expt. 9, the spatial rainfall distributions under the $R_{assumed}$ scenario were generated from the three raingauges by using IDW, KG, and TSP methods, respectively. These spatial rainfall distributions are presented in Figure 7.12.



(a)

(b)

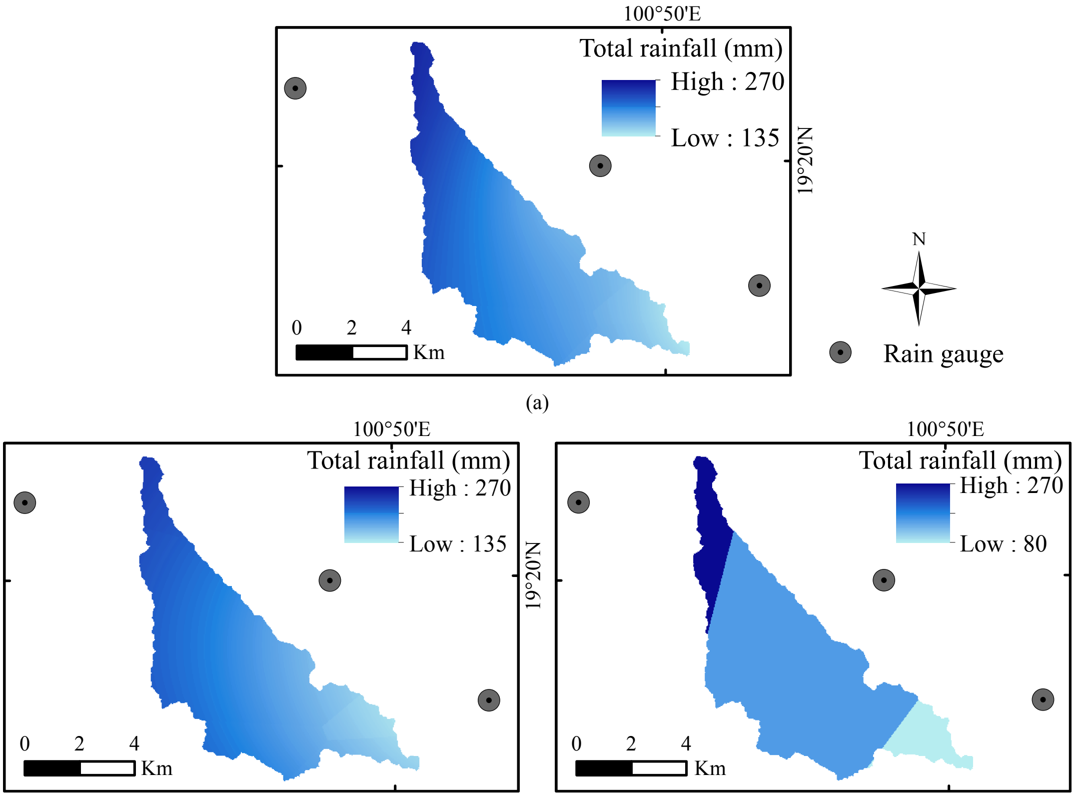
(c)

(d)

(f)

Figure 7.11: Spatial distributions of the three day accumulated rainfall between 25 and 27 June 2011 in the Nam Hui watershed (N2): (a) for Expt. 2 derived from radar-rainfall measurements and (b) for Expt. 3 derived by the IDW, (c) for Expt. 4 derived by the KG, (d) for Expt. 5 derived by the TSP methods from the three raingauges, and (f) for Expt. 6 derived by the IDW method from the two raingauges.

In the Expt. 1, the PCA method was performed on the covariance matrix of the selected six physiographic variables. All PCs of the Expt. 1, their principal component coefficients, and their variance are shown in Table 7.7. It reveals that the first four PCs together accounted for 95% of the total variance in the original data set.



(a)

(b)

(c)

Figure 7.12: Spatial rainfall distributions of the $R_{assumed}$ scenario in the Nam Hui watershed (N2): (a) for Expt. 7 derived by the IDW, (b) for Expt. 8 derived by the KG, and (c) for Expt. 9 derived by the TSP methods from the three raingauges.

Table 7.7: All principal components and their variance for the Expt. 1 (physiographic variables)

Variables	PC1	PC2	PC3	PC4	PC5	PC6
Dd	0.11	**0.98**	0.03	-0.04	0.08	0.11
Rbuffer	**0.55**	-0.08	-0.26	-0.08	**0.72**	-0.32
Surface slope	0.13	-0.07	**0.95**	-0.17	0.22	-0.02
Elevation	**0.55**	0.05	0.07	0.07	**-0.62**	**-0.55**
TWI	0.04	0.02	0.14	**0.98**	0.14	0.04
FL	**0.61**	-0.14	-0.04	-0.02	-0.16	**0.76**
Variance [%]	54.40	20.69	13.82	6.16	4.30	0.64
Cumulative variance [%]	54.40	75.09	88.90	95.06	99.36	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Moreover, PCA results of the Expt. 1. are visually interpreted by using a three-dimensional (3D) plot (Figure 7.13a). The axes of this plot are the first three principal components: PC1, PC2, and PC3. The straight lines represent variables. The lengths of the straight lines indicate how each variable contributes to the first three PCs. The angles among the straight lines indicate how each variable correlates to others. If two variables strongly correlated, the angle between their lines is small. As it can be seen in Figure 7.13a and Table 7.7, PC1 was strongly correlated with FL, elevation, and Rbuffer and accounted for more than 54.4% of the total variance. PC2, PC3, and PC4 described 20.7%, 13.8%, and 6.2%, respectively, of the total variance. PC2 was mainly correlated with Dd and PC3 was mainly described by the surface slope. PC4 was dominantly correlated with TWI.

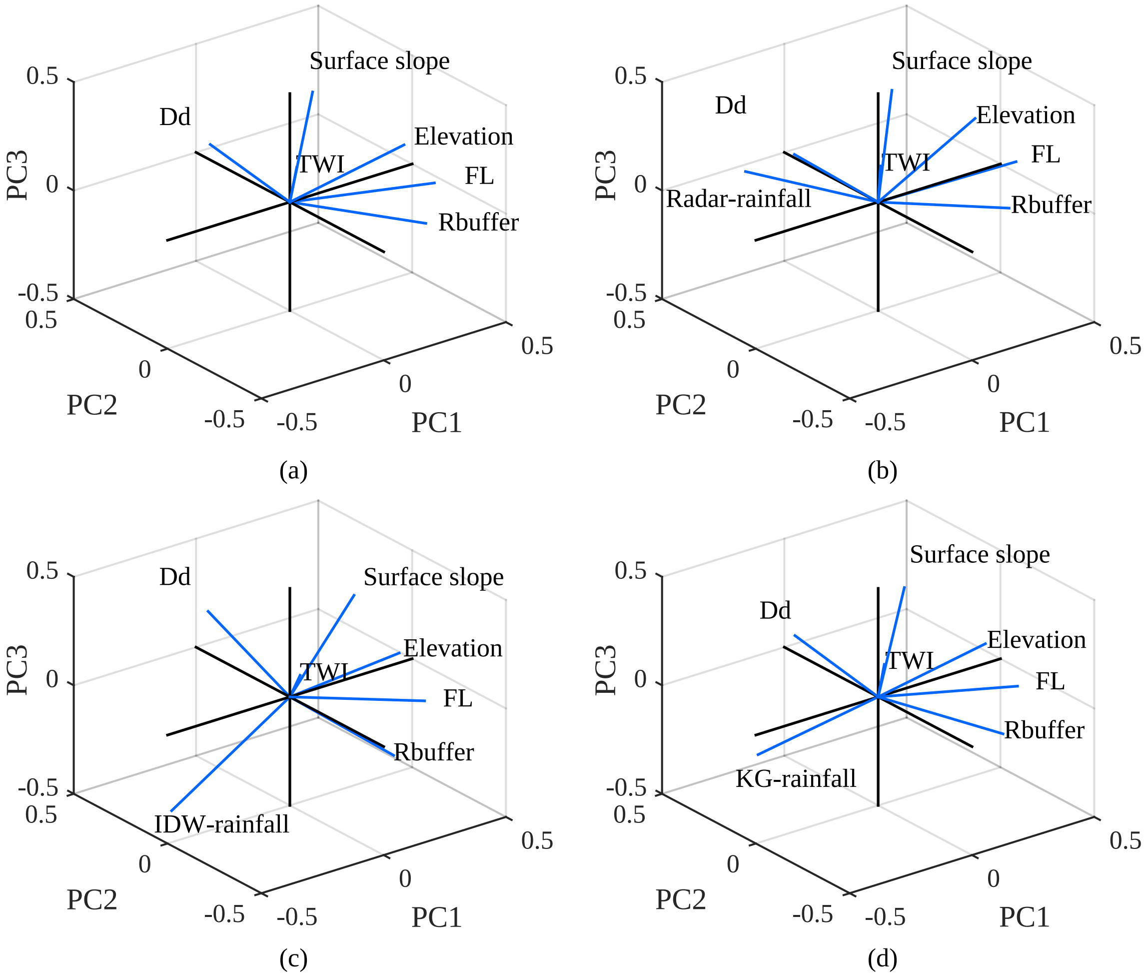


Figure 7.13: Three-dimensional plots of the first three principal components: (a) for Expt. 1, (b) for Expt. 2, (c) for Expt. 3, and (d) for Expt. 4.

Afterwards, flash flood potential indices (FFPI) were calculated by using Equation (6.3). For example, the FFPI values of the Expt. 1 can be expressed as FFPI = 0.57PC1+0.22PC2+0.15PC3+0.06PC4. The spatial distribution of FFPI values is discussed later in this section.

In the Expt. 2, the combination of the six physiographic variables and the spatial radar-rainfall distribution was summarized by PCA. PCA results of this Expt. 2 are shown in Table 7.8 and Figure 7.13b. The first four PCs explained about 91% of the total variance. PC1 represented 55.8% of the total variance and was strongly correlated with FL and Rbuffer. For PC1, it was found that the spatial radar-rainfall distribution did not correlate with the other variables on the PC1 axis. PC2 and PC3 accounted for 17.7% and 12.2%, respectively, of the total variance. The main contributors of PC2 and PC3 for the Expt. 1 and Expt. 2 were identical. PC4 was significantly correlated with the radar-rainfall distribution and accounted for 5.5% of the total variance.

Table 7.8: All principal components and their variance for the Expt. 2 (physiographic variables and radar-rainfall spatial distribution under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.08	**0.98**	-0.05	-0.13	0.05	0.03	0.10
Rbuffer	**0.50**	-0.05	-0.18	-0.01	-0.03	**0.80**	-0.25
Surface slope	0.09	-0.01	**0.93**	-0.33	0.04	0.14	-0.03
Elevation	0.48	0.11	0.17	0.42	-0.27	-0.43	**-0.54**
TWI	0.03	0.03	0.14	**0.53**	**0.83**	0.06	0.03
FL	**0.55**	-0.09	0.05	0.18	-0.17	-0.10	**0.78**
Radar-rainfall	-0.44	0.13	0.24	**0.61**	-0.45	0.37	0.12
Variance [%]	55.78	17.74	12.23	5.51	5.02	3.24	0.47
Cumulative variance [%]	55.78	73.52	85.75	91.27	96.29	99.53	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

In the Expt. 3, the physiographic and climate variables were transformed to a new set of components by using PCA. This was similar to the Expt. 2. However, the difference between both experiments was that the spatial rainfall distribution for Expt. 3 was created by using the IDW interpolation of the three raingauges instead of using radar-rainfall observations. The relationship between the variables of the Expt. 3 was described by the first four PCs, which explained about 92% of the total variance as shown in Table A.3 (in Appendix A.3.2). Figure 7.13c shows correlations of the variables in a 3D plot for the Expt. 3. PC1, PC2, PC3, and PC4 represented 51.4%, 17.6%, 15.2%, and 7.8%, respectively, of the total variance. Compared to Expt. 1, FL and Dd mainly contributed to PC1 and PC2, respectively. PC3 and PC4 were loaded on surface slope.

In the Expt. 4, PCA was run on the physiographic variables and spatial rainfall distribution interpolated by using the KG method from the three raingauges. Scanning Table A.4 (in Appendix A.3.2), it can be seen that the first four PCs used to construct the FFPI made up about 92% of the total variance. Figure 7.13d shows correlations of the variables in a 3D plot for the Expt. 4. PC1, which indicated 55.7% of the total variance, was mainly related to FL and elevation. Dd significantly contributed in PC2, which accounted for 18.2% of the total variance. PC3 and PC4, which represented about 12.3% and 5.4% of the total variance, were mainly correlated to surface slope and TWI, respectively. In short, the significant contributors of the first four PCs in the Expt. 4 were compatible with the Expt. 1 and Expt. 2. These main contributors were FL, elevation, Rbuffer, Dd, and surface slope.

In the Expt. 5, PCA identified the pattern of the physiographic variables and spatial rainfall distribution interpolated by using the TSP method from the three raingauges. This pattern can be interpreted by plotting principal component scores in a plot shown in Figure 7.14. The plot for the Expt. 5 represents the principal component scores, the so-called “data cloud”, on the first and second principal component axes. The data cloud represents the 325 900 points of the data set. To capture the maximum variability in the data cloud, PCA searched for a linear combination of variables that was significantly fit with the data cloud in the least squares sense. Put differently, PCA minimized the sum of squared distance between the original data set and linear combinations (axes) of the variables and projected the data set to the axes (Johnson and Wichern, 2002). The axis, which accounted for the most variance in the data and had least-squared distance, was then defined as PC1.

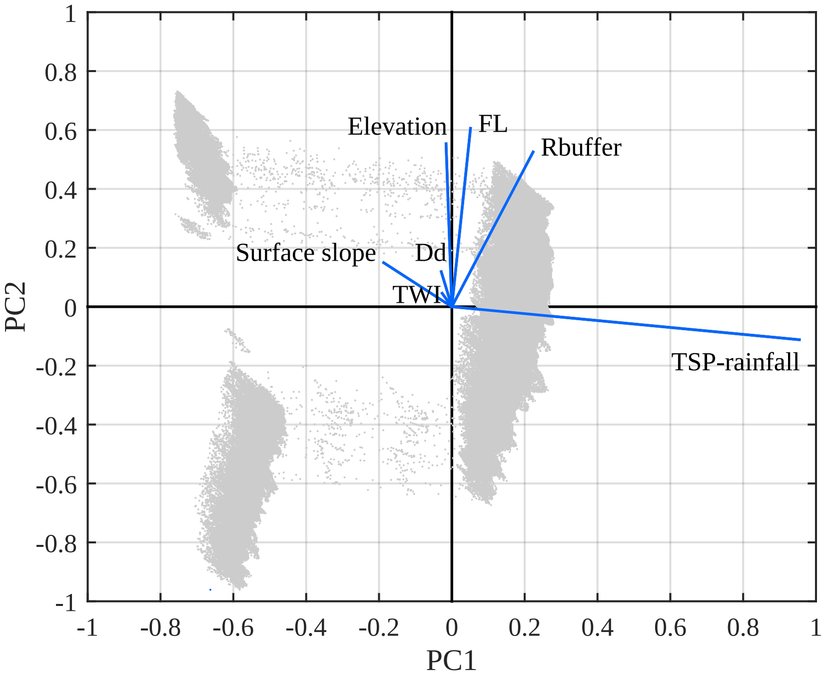


Figure 7.14: Plot of the principal component scores on the first and second principal component axes for the Expt. 5.

For example, the PC1 and PC2 axes of the Expt. 5, which run through the centroid of the data cloud, accounted for 37.0% and 35.6% of the total variance (Table 7.9). In PC1, the spatial TSP-rainfall played an important role. It can be also seen from Figure 7.14 that a straight line of the spatial TSP-rainfall variable originates from the centroid of the data cloud and reaches an end at a pair of coordinates (0.95, -0.11) on the PC1 and PC2 axes. This pair of the coordinates (0.95, -0.11) indicates the coefficients of the spatial TSP-rainfall variable corresponding to the first two linear components (PC1 and PC2), respectively. The positive coefficient shows that the variable had a positive correlation to PC1, whereas the negative coefficient shows that the variable had a negative correlation to PC2. Moreover, the main contributors of PC2 were FL, elevation, and Rbuffer. The coefficients of the FL, elevation, and Rbuffer variables were 0.61, 0.55, 0.53, respectively, as shown in Table 7.9. The two following components (PC3 and PC4) represented about 13.5% and 7.4%, respectively, of the total variance. The mainly contributing variables of PC3 and PC4 were Dd and surface slope, respectively.

The most interesting aspect of the plot (Figure 7.14) is that the data cloud is separated into three main groups. This is a consequence of the interpolated TSP-rainfall distribution from the three raingauges. The TSP interpolation method computes every point rainfall value within the watershed, which is equal to the observed rainfall at the nearest gauge. In this case, the three rainfall values were constantly distributed throughout the watershed. Consequently, the data cloud was broken down into three main parts. This data distribution phenomenon may lead to a mistaken interpretation because PCA assumes that relationships among the original variables are predominantly linear (Jolliffe, 2002). In other words, if the data distribution lies on differentiable manifolds or is prominently nonlinear, the problem to summarize the original data set may occur because PCA is sensitive to the presence of outliers, which are significantly apart from the rest (Locantore et al., 1999).

Table 7.9: All principal components and their variance for the Expt. 5 (physiographic variables and TSP-rainfall spatial distribution derived from the three raingauges under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	-0.03	0.12	**0.98**	0.08	-0.04	0.07	0.11
Rbuffer	0.22	**0.53**	-0.06	-0.10	-0.05	**0.72**	-0.38
Surface slope	-0.19	0.15	-0.11	**0.95**	-0.14	0.06	-0.01
Elevation	-0.02	**0.55**	0.04	-0.04	0.05	**-0.66**	**-0.50**
TWI	-0.03	0.05	0.01	0.13	**0.99**	0.08	0.04
FL	0.05	**0.61**	-0.15	-0.10	-0.03	-0.08	**0.77**
TSP-rainfall	**0.95**	-0.11	0.03	0.22	0.01	-0.16	0.04
Variance [%]	36.95	35.60	13.52	7.39	4.03	2.14	0.38
Cumulative variance [%]	36.95	72.54	86.06	93.45	97.48	99.62	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

PCA results for the last four experiments (Expt. 6–9) are shown in Table A.6 - A.9 (in Appendix A.3.2). These tables present all principal components and their variance. In addition, the tables show component coefficients of the PCs, which represent the correlations of variables with these PCs. These results can be interpreted in similar ways as the above explanation of the first five experiments (Expt. 1–5). Furthermore, PCA results for all the nine experiments are summarized in Table 7.10. This table presents significant variables that contribute to the first four PCs of the nine experiments.

Table 7.10: Summary of mainly contributing variables for the first four PCs of the nine experiments

Experiment	Mainly contributing variables for			
	PC1	PC2	PC3	PC4
Expt.1	Flow length	Drainage density	Surface slope	TWI
Expt.2	Flow length	Drainage density	Surface slope	Radar-rainfall
Expt.3	Flow length	Drainage density	Surface slope	Surface slope
Expt.4	Flow length	Drainage density	Surface slope	TWI
Expt.5	TSP-rainfall	Flow length	Drainage density	Surface slope
Expt.6	IDW-rainfall	Drainage density	Surface slope	Rbuffer
Expt.7	Flow length	Drainage density	Surface slope	TWI
Expt.8	KG-rainfall	Drainage density	Surface slope	Rbuffer
Expt.9	Flow length	Drainage density	Surface slope	TSP-rainfall

Note: Rbuffer (distance from stream), TWI (topographic wetness index),
IDW (inverse distance weighting), KG (kriging), and TSP (Thiessen polygon)

As aforementioned, flash flood potential indices were calculated by using Equation (6.3). Spatial distributions of FFPI values for all the experiments are shown in Figure 7.15. The spatial pattern of FFPI values for the Expt. 1 is presented in Figure 7.15a. Its high values indicating high flash flood potential occur in downstream areas, while low FFPI values are located in upstream areas.

Furthermore, the spatial FFPI distributions of the Expt. 2, 3, 4, 7, and 9 are shown in Figure 7.15 and compared to the FFPI of the Expt. 1. Higher FFPI values were delineated in downstream areas, whereas upstream areas are characterized by lower FFPI values. On the other hand, higher FFPI values were located in upstream areas, and lower FFPI values were found in downstream areas for the Expt. 6 and Expt. 8 as

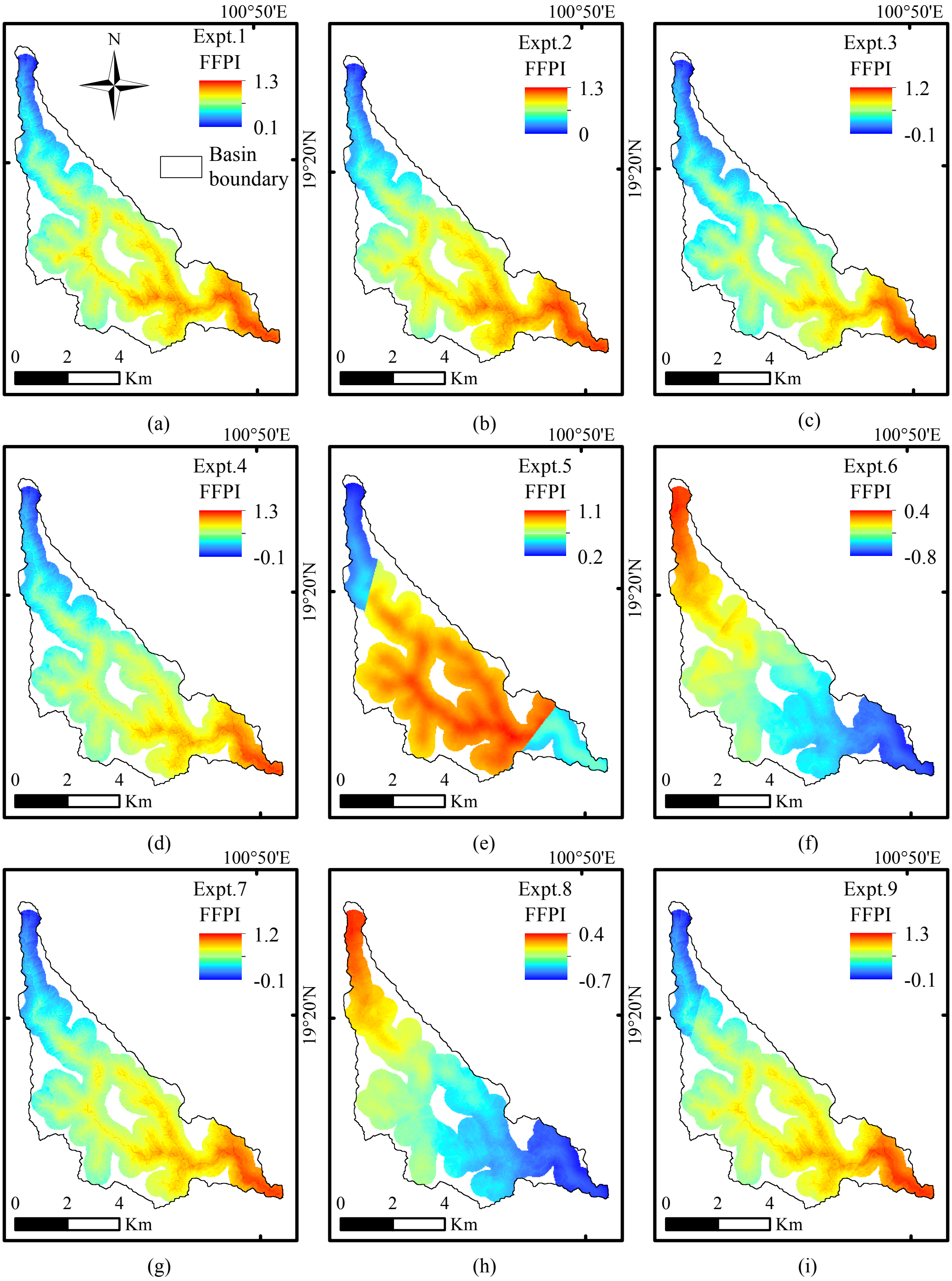


Figure 7.15: Maps of spatial FFPI distributions based on the physiographic variables of the Nam Hui watershed (N2): (a) excluded rainfall, (b) included radar-rainfall distribution under the R250611 scenario, (c) included IDW-rainfall, (d) included KG-rainfall, and (e) included TSP-rainfall distributions derived from the three raingauges under the R250611 scenario, (f) included IDW-rainfall distribution derived from the two raingauges under the R250611 scenario (g) included IDW-rainfall, (h) included KG-rainfall, and (i) included TSP-rainfall distributions derived from the three raingauges under the $R_{assumed}$ scenario.

shown in Figure 7.15f and 7.15h, respectively. Moreover, some of FFPI values were negative. This can be explained by the results of the PCA analysis as shown in Table A.6 and A.8 (in Appendix A.3.2). What stands out in these tables is the non-correlation among the significant variables. For example, PC1 had a highly positive correlation with the spatial IDW-rainfall variable but a highly negative correlation with FL for the Expt. 6 (see Table A.6). According to the spatial rainfall distributions (IDW-rainfall and KG-rainfall) of the Expt. 6 and Expt. 8, a higher amount of rainfall was spatially distributed in the upstream areas of the watershed as previously shown in Figure 7.11f and 7.12b, respectively. The spatial distribution of the FL was consistent with the spatial rainfall distribution that higher values were spatially distributed in the upstream areas of the watershed. However, the FL variable was considered as an inverse factor that reduces the degree of flash flood potential (as discussed in Section 6.3.1.2). Consequently, lower values of FL being shorter to the basin outlet reflect the higher flood potential at the site. As a result, the strong uncorrelations between the two significant variables of PC1 for the Expt. 6 and Expt. 8 produced the negative FFPI values. For these two experiments, it was found that the PCA method cannot represent dynamics in overland flow routing.

For the Expt. 5, higher FFPI values were delineated in the middle part of the watershed as shown in Figure 7.15e. The distribution of the FFPI values for this experiment accounted for the effect of the spatial rainfall distribution interpolated from the three raingauges by using the Thiessen polygon method. This TSP-rainfall distribution, which was the main contributor of PC1 as shown in Table 7.10, interfered the variability of the whole data pattern. As a result, there were three dominant parts of the spatial FFPI distribution as presented in Figure 7.15e.

According to the sensitivity analysis of the spatial index-based approach to different spatial rainfall distributions, the present study noted that results of this spatial index-based approach were very sensitive to patterns of spatial rainfall distributions. Thus, the spatial rainfall distributions were omitted from the spatial index-based approach at the present stage of this research.

To summarize, six physiographic variables (i.e., Dd, Rbuffer, surface slope, elevation, TWI, and FL) were used to construct FFPI. Once FFPI values of the five sub-basins in the upper Nan River basin were computed, these values needed to classify in order to determine degrees of flash flood hazard. Results of the FFPI classification are presented later in Section 7.4.2.

7.4 Flash flood hazard mapping

This section presents flash flood hazard mapping based on the results obtained by the integrated modeling approach (as described in Section 7.2) and by the spatial index-based approach (as described in Section 7.3).

7.4.1 Flash flood hazard maps based on the integrated modeling approach

The MEADFLOW hydrodynamic model has the capability to produce 2D maps of the maximum flood inundation and velocity in the river and floodplain. These 2D maps were used to compute flood hazard ratings by using the Equation 6.4. The flood hazard ratings were classified into five levels of flood severities: extreme, significant, moderate, low, and very low. This classification was established on the basis of potential flash flood hazards to people as a function of water depth and velocity (Ramsbottom et al., 2006).

As an example, a flash flood hazard map for downstream areas of the Nam Rim watershed (N1) based on the regional rainfall scenario of 10-year return period is shown in Figure 7.16.

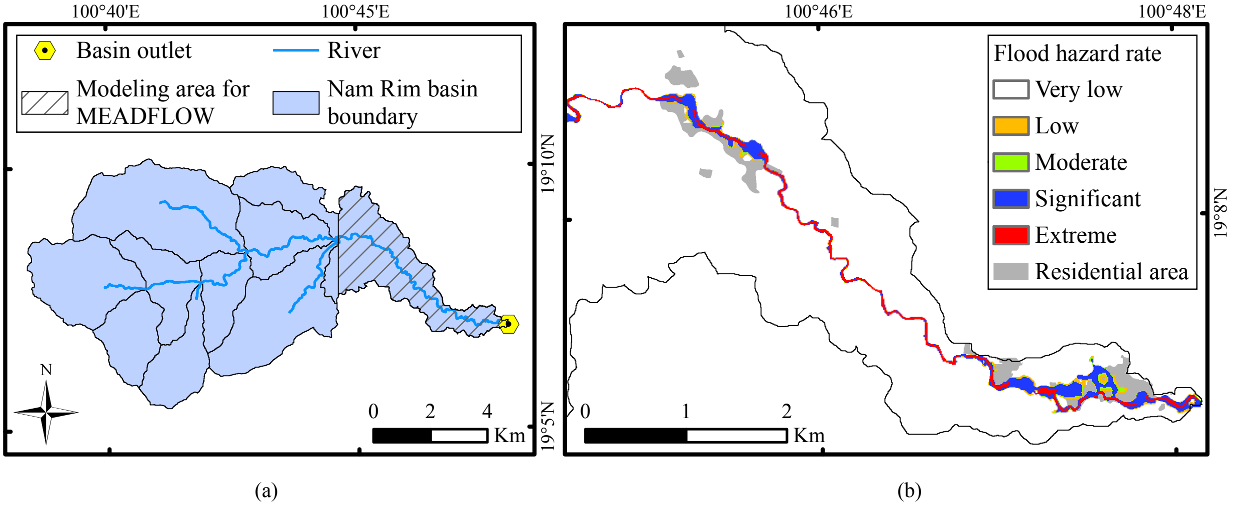


Figure 7.16: Flood hazard map of the Nam Rim watershed based (N1) on the regional rainfall scenario of RTr10.

In sum, flash flood hazard maps based on the integrated modeling approach were produced for the five sub-basins of the upper Nan River basin. These flash flood hazard maps were later compared with those produced by the spatial index-based approach (as described in Section 7.4.2).

7.4.2 Flash flood hazard maps based on the spatial index-based approach

7.4.2.1 Classification of flash flood potential indices

To analysis and classify FFPI based on the PCA method, spatial FFPI grids of the five sub-basins in the upper Nan River basin were pooled together. The values of the FFPI grids were normalized and the distribution of them was compared with the theoretical distribution candidates: generalized extreme value (GEV), Gumbel and lognormal distributions. Figure 7.17 shows the comparison between the cumulative distribution curve of the FFPI values and the curves of these candidates. On the basis of the Bayesian information criterion developed by Schwarz (1978), the distribution curve of the FFPI values was best fitted to the GEV distribution.

The histogram shown in Figure 7.18a illustrates the probability distribution of the FFPI values based on the normalized FFPI values of the five sub-basins in the upper Nan River basin and compares with the GEV probability density function. The 98th percentile of the GEV probability distribution was considered suitable to be used as a lowest classifier threshold for classifying flash flood hazards in the five sub-basins. In order to avoid confusing use of this lowest classifier threshold, the percentile does not refer to an annual nonexceedance probability for a series of events (i.e., the 98th percentile means a flood with a return period of 50 years). However, the percentile refers to the position of an individual on a given probability density distribution. For example, the FFPI value at the position of the 98th percentile on the GEV probability distribution is 0.9 (Figure 7.18). Afterwards, the lowest classifier threshold was used to determine a lower boundary for omitting areas in which the occurrence of flash floods is extremely rare from the analysis of the index classification.

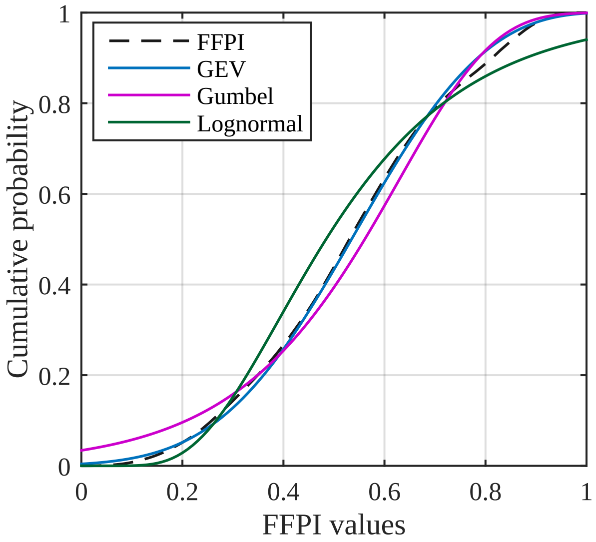


Figure 7.17: Cumulative density curves of FFPI based on data of all the five sub-basins and theoretical distribution candidates.

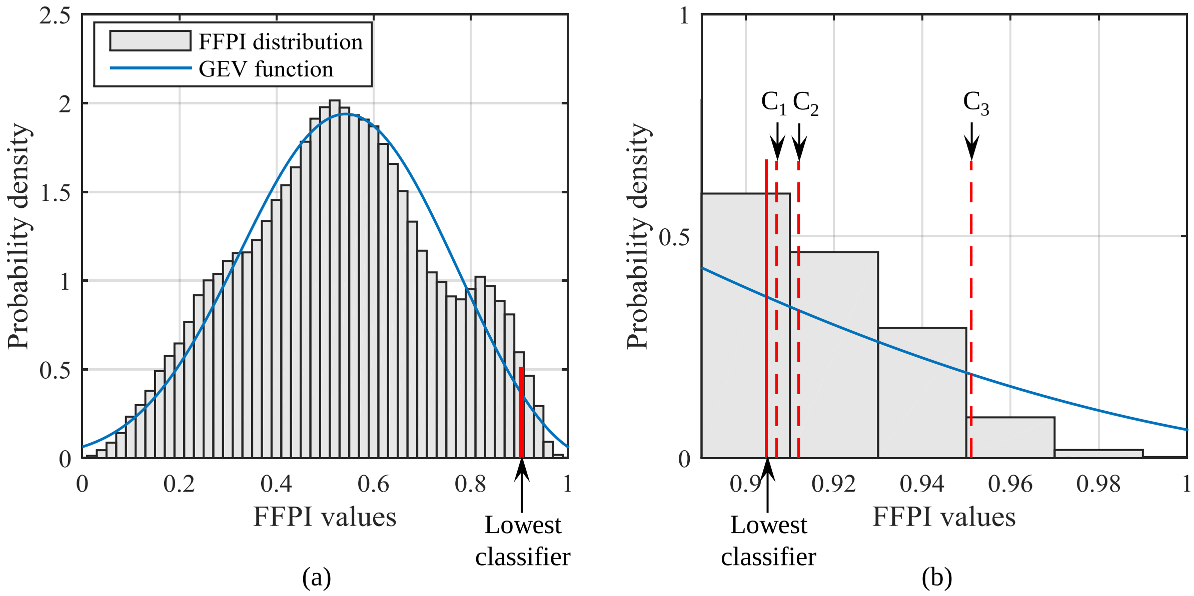


Figure 7.18: FFPI distribution based on data of all the five watersheds and their classifiers: (a) distribution of FFPI values fitted to generalized extreme value (GEV) probability density function and a lowest classifier and (b) right tail of FFPI distribution and three classifiers (C_1, C_2, and C_3) for identifying low to extreme hazard areas.

According to information about flash flood occurrence collected from the field investigations of this dissertation, it was found that flash floods did not yet occur in areas where indicated by FFPI values below the lowest classifier threshold (the 98th percentile of the GEV probability distribution). Therefore, these areas were identified by a very low rate of flash flood hazards. However, the FFPI values above the 98th percentile of the GEV probability distribution were classified into four flood hazard categories. Their corresponding areas were identified with extreme, significant, moderate and low hazard rates.

To search out a set of classifiers, which was suitable for classifying a range between the lower and upper boundaries, 100 000-sample sets of three random classifiers were generated. For example, three random classifiers (C_1, C_2, and C_3) were any three percentiles within the the 98th and 100th percentiles of the

GEV probability distribution (Figure 7.18b). The three random classifiers of each set were used to group the FFPI values, which were in the range, into four flood hazard categories. Once the FFPI values were grouped, their corresponding areas were labeled by four flood hazard rates: low, moderate, significant, and extreme hazards. As the 100 000-sample sets of three classifiers were used to classify the flood hazard categories, 100 000 flash flood hazard maps were created. These maps were compared with flash flood hazard maps of the five selected watersheds that were produced by the integrated modeling approach. Moreover, they were evaluated by three statistical measures: agreement ratio, *F*-score, and *RMSE* (see Section 3.4 and Appendix A.2.1). Results of these statistical measures for the regional rainfall scenario of RTr10 were plotted in a 3D scatter plot (Figure 7.19). The X-axis, Y-axis, and Z-axis of the 3D scatter plot represent agreement ratio, *RMSE*, and *F*-score, respectively. In the plot, there are 100 000 points representing 100 000-sample sets of classifiers. The suitable set of the classifiers, which was highlighted by the red sphere in the 3D scatter plot, was chosen on the basis of optimizing the three statistical measures. For the agreement ratio and *F*-score, two maps perfectly agree when their values equal to one. For the *RMSE*, the value closer to zero indicates the better agreement between two maps.

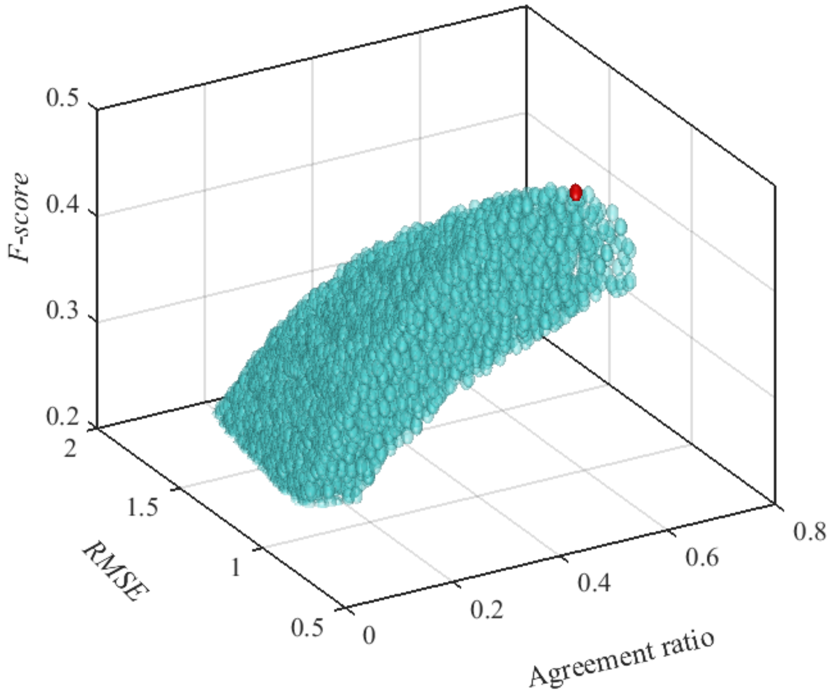


Figure 7.19: 3D scatter plot of three statistical measures (agreement ratio, *RMSE*, and *F*-score) evaluated flash flood hazard maps for the regional rainfall scenario of RTr10 produced by the spatial index-based approach based on 100 000-sample sets and produced by the integrated modeling approach and the red sphere represented the result of the suitable set for classifying FFPI values based on the three statistical measures.

The 100 000 flash flood hazard maps of the five watersheds were compared with the flash flood hazard maps of the ones under the regional rainfall scenario of RTr10 produced by the integrated modeling approach. In order to evaluate the comparison, the three statistical measures were used. On the basis of the three statistical measures, it was found that a set of three classifiers, which were 98.02th, 98.04th, and 99.26th percentiles of the GEV probability distribution fitted to the FFPI distribution of the five watersheds, was suitable for classifying the FFPI values into the four flood hazard categories: low, moderate, significant, and extreme for the regional rainfall scenario of RTr10. The agreement ratio and average *F*-score measures were 0.56 and 0.48, respectively. The *RMSE* measure evaluated the set of the three classifiers by 0.92 with comparing to the flash flood hazard maps of the five watersheds under the regional rainfall scenario of RTr10 produced by the integrated modeling approach.

For instance, the set of three classifiers was used to categorize the FFPI values of the Nam Rim watershed (N1) into four flood hazard categories: low, moderate, significant, and extreme. The areas indicating FFPI values below the 98th percentile of the GEV probability distribution were identified to have a very low rate of flash flood hazards. As a result, these five flood hazard categories in total were used to delineate a flash flood hazard map under the regional rainfall scenario of RTr10. Figure 7.20a shows the flash flood hazard map produced by the integrated modeling approach and Figure 7.20c shows a closer view of the flash flood hazard map based on the integrated modeling approach. Moreover, Figure 7.20b shows the flash flood hazard map produced by using the spatial index-based approach with the set of three classifiers based on the agreement with the flash flood hazard map of the integrated modeling approach. A closer view of the flash flood hazard map based on the spatial index-based approach is displayed in Figure 7.20d. As can be seen from Figures 7.20c and 7.20d, extreme flood hazard areas indicated by the spatial index-based approach are in good agreement with those indicated by the integrated modeling approach. However, significant flood hazard areas identified by the spatial index-based approach are larger than those obtained from the integrated modeling approach.

In addition to visual representation of the flash flood hazard maps produced by the both approaches as shown in Figure 7.20, a statistical comparison was performed by using the three statistical measures. The agreement ratio indicated overall effectiveness of the classifiers with 0.56 and the average *F*-score indicated with 0.51. Moreover, classified areas of the both flash flood hazard maps were evaluated by the *RMSE* with -0.10. According to the three statistical measures, the flash flood hazard map produced by using the spatial index-based approach with the set of three classifiers had a good agreement with the one produced by the integrated modeling approach. For flash flood hazard maps of the other watersheds, they are displayed in Appendix A.3.3.

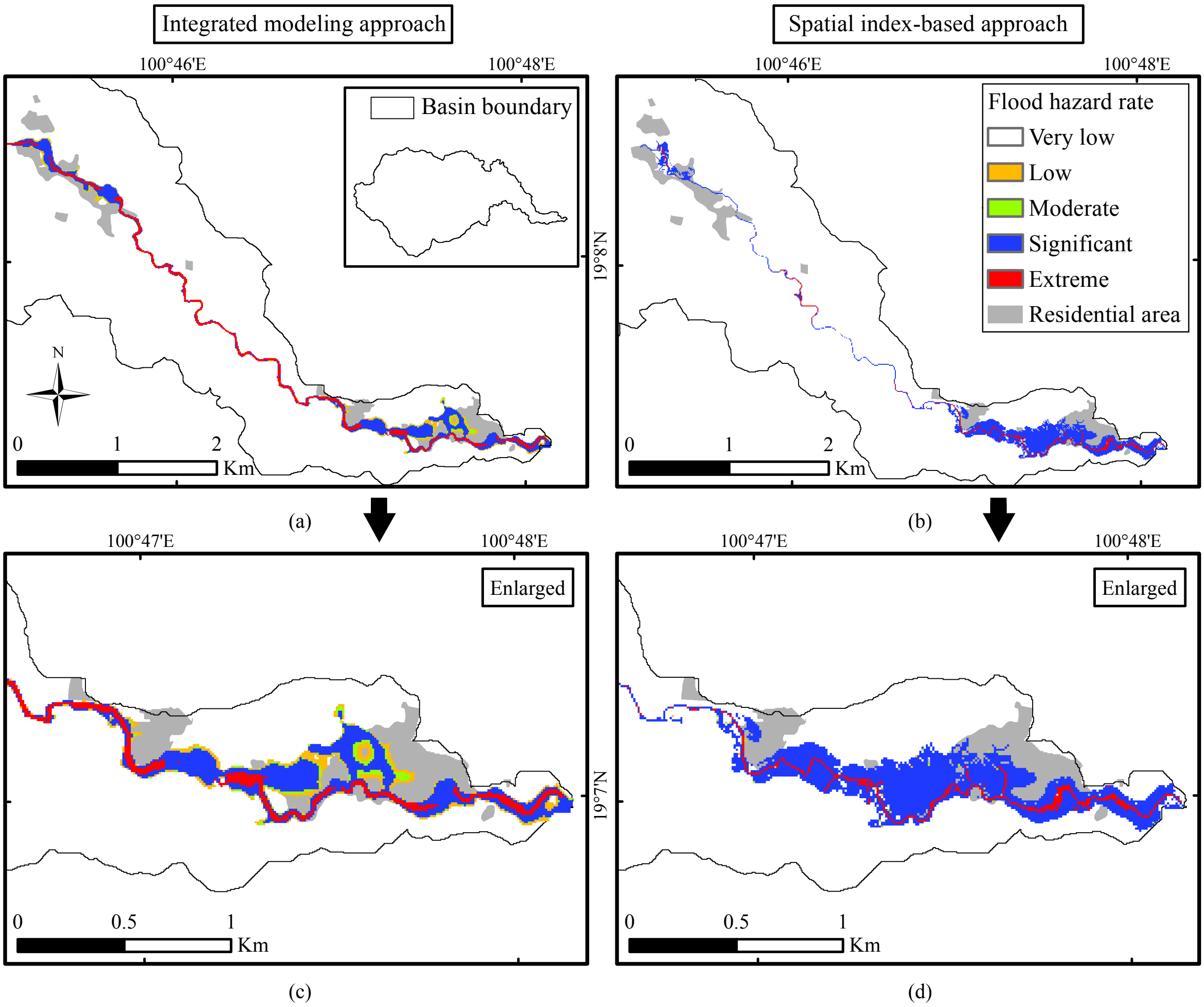


Figure 7.20: Flash flood hazard maps of the Nam Rim watershed (N1) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach, (b) produced by the spatial index-based approach, (c) produced by the integrated modeling approach (with a closer view), and (d) produced by the spatial index-based approach (with a closer view).

7.4.2.2 Comparison of flash flood hazard maps under different rainfall scenarios

The previous section presents the classification of FFPI based on the analysis of all five watersheds together in order to discover regional classifiers to define flash flood hazard zones. Conversely, this section presents the classification of FFPI that was based on an individual watershed. As a result of the FFPI classification, classifiers were obtained by comparing their corresponding hazard areas with flash flash hazard areas computed by using the integrated modeling approach.

The Nam Rim watershed (N1) was chosen to demonstrate and investigate the use of FFPI classification for applying to a single watershed. Prior to the FFPI classification, it was necessary to identify a probability distribution that can represent an FFPI distribution of a given watershed. Therefore, the FFPI distribution of the Nam Rim watershed was compared with commonly used distributions – namely, GEV, Gumbel, and lognormal probability distributions. It was found that the GEV probability distribution was best fitted to the FFPI distribution of the Nam Rim watershed on the basis of the Bayesian information criterion (Schwarz, 1978). This was compared to the analysis of the FFPI distribution based on all the five watersheds together.

Figure 7.21a represents the probability distribution of the FFPI values based on the data of the Nam Rim watershed (N1) and the GEV probability density function that was the best-fit function to the FFPI distribution. The lowest classifier threshold for classifying flash flood hazards in the Nam Rim watershed was the 98th percentile of the GEV probability distribution that was in accordance with the one evaluated on the basis of all the five watersheds together. The lowest classifier threshold for the Nam Rim watershed was obtained by comparing areas of low to extreme hazards with inundated areas simulated by the integrated modeling approach under different rainfall scenarios – namely, R130D6, R170D6, and R200D6 scenarios. For example, the R130D6 scenario means an event where a rainfall amount of 130 mm was assumed to be uniformly distributed over a given watershed for a duration of six hours.

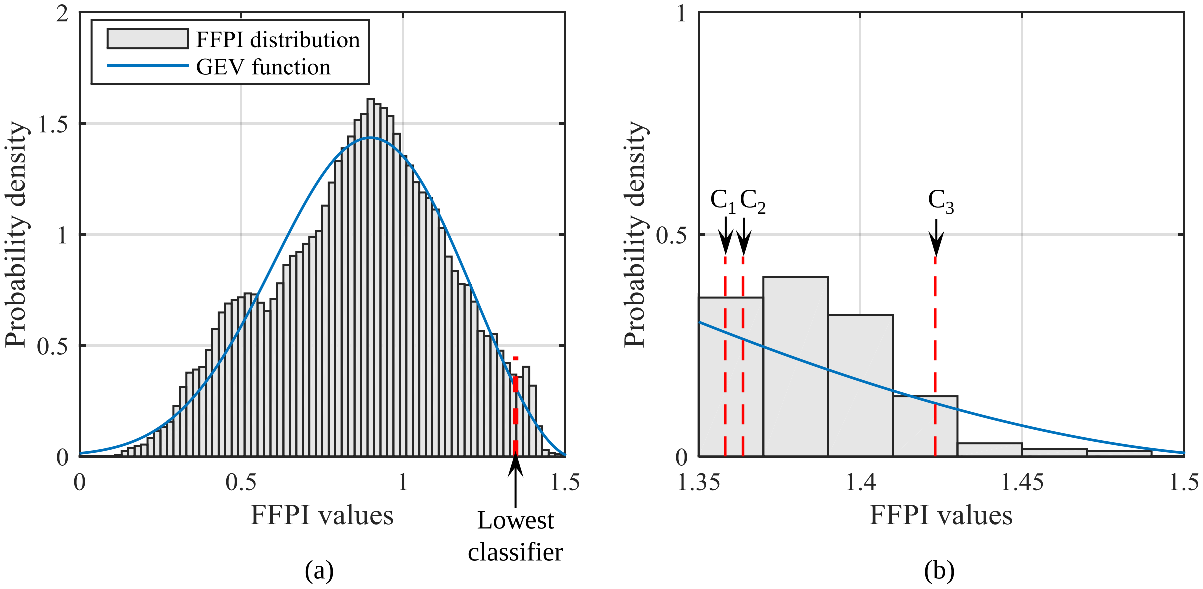


Figure 7.21: FFPI distribution based on data of the Nam Rim watershed (N1) and their classifiers: (a) distribution of FFPI values fitted to generalized extreme value (GEV) probability density function and a lowest classifier and (b) distribution of FFPI values greater than the lowest classifier and three classifiers (C_1, C_2, and C_3) for identifying low to extreme hazard areas.

To evaluate classifiers, the FFPI values of the Nam Rim watershed were grouped into five categories and their corresponding areas were labeled by five flood hazard rates: very low, low, moderate, significant, and extreme. As previously mentioned, the lowest classifier was used to separate FFPI values of a given watershed into two zones that were very low and low to extreme hazards. The lowest classifier was obtained by comparing the corresponding areas of low to extreme hazards with flood extent simulated by the integrated modeling approach. To categorize the areas of low to extreme hazards, the FFPI values that were greater than the lowest classifier were grouped into the four flood hazard rates by a set of three classifiers (C_1, C_2, and C_3) as shown in Figure 7.21b. Afterwards, low, moderate, significant, and extreme hazards were assigned to their corresponding areas.

As mentioned above, there were three different rainfall scenarios (R130D6, R170D6, and R200D6 scenarios) used as input to the integrated models in order to create flash flood hazard maps. These three maps were compared with hazard areas delineated by three sets of three classifiers and were used to evaluate these three sets of the three classifiers. Table 7.11 presents the three sets of the three classifiers used to group the FFPI values greater than the lowest classifier into the four flood hazard rates. For instance, the C_1, C_2, and C_3 classifiers were the 98.30th, 98.45th, and 99.57th percentiles of the GEV probability distribution, respectively, used to group the FFPI values and delineate flash flood hazard areas for the R130D6 rainfall scenario. These flash flood hazard areas based on the FFPI classification were compared with the flash flood hazard map of the R130D6 rainfall scenario produced by the integrated models (Figure 7.22). The comparison results of the R130D6 rainfall scenario are in fair agreement indicated by an agreement ratio of 0.52.

Table 7.11: Classifiers used to classify FFPI values based on the Nam Rim watershed for different rainfall scenarios

Scenario	Classifiers [-th percentile of GEV probability distribution] C_1	C_2	C_3	Agreement ratio* [-]
R130D6	98.30	98.45	99.57	0.52
R170D6	98.27	98.43	99.50	0.52
R200D6	98.25	98.40	99.45	0.51

Note: *Congalton (1991)

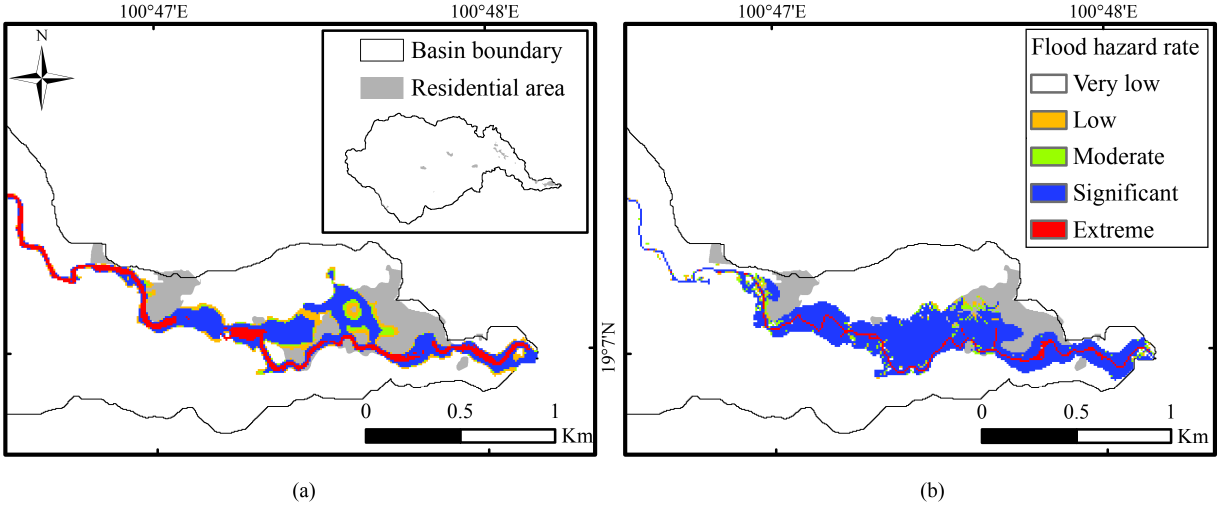
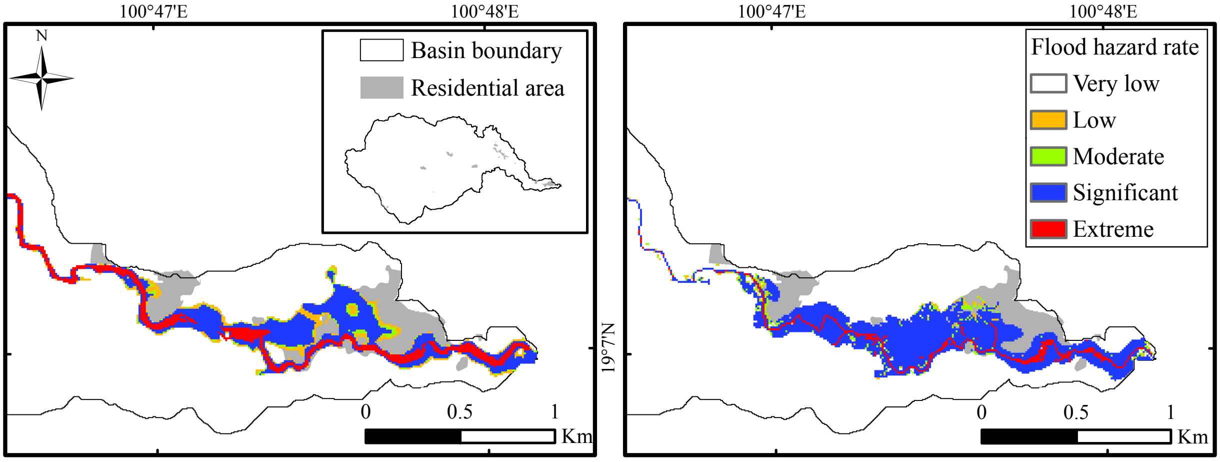


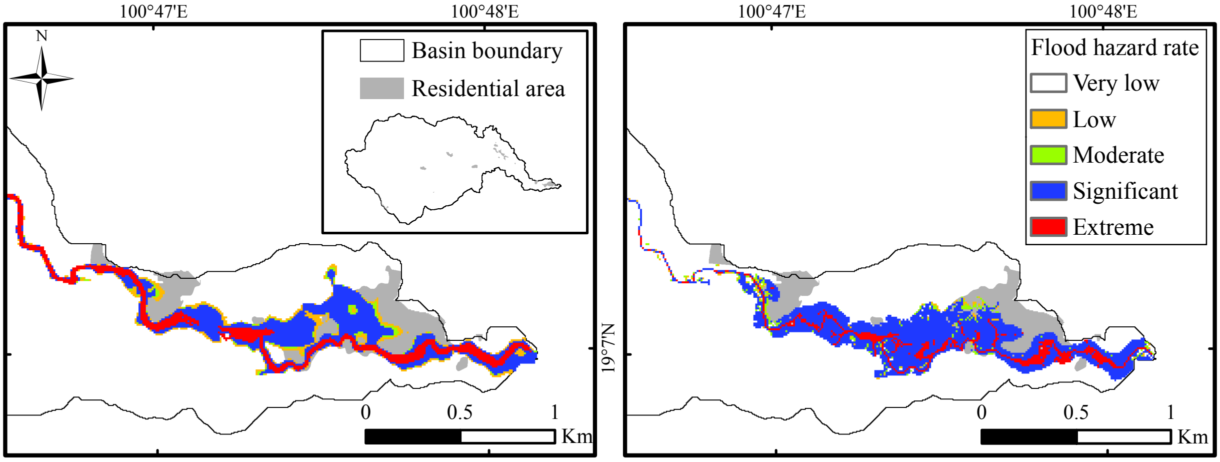
Figure 7.22: Flash flood hazard maps of the Nam Rim watershed (N1) under the R130D6 scenario: (a) produced by the integrated modeling approach, (b) produced by the spatial index-based approach.

Moreover, Figures 7.23 and 7.24 show flash flood hazard maps of the R170D6 and R200D6 scenarios, respectively, for the Nam Rim watershed (N1) produced the integrated modeling approach and the spatial index-based approach. As it can be seen from the figures, flash flood hazard maps produced by the spatial index-based approach are in fair agreement with those produced by the integrated modeling approach. The results appear to be well supported by the areas identified as the extreme level of flash flood hazards by the both approaches because they are well good consistencies in their comparisons.



(a) (b)

Figure 7.23: Flash flood hazard maps of the Nam Rim watershed (N1) under the R170D6 scenario: (a) produced by the integrated modeling approach, (b) produced by the spatial index-based approach.



(a) (b)

Figure 7.24: Flash flood hazard maps of the Nam Rim watershed (N1) under the R200D6 scenario: (a) produced by the integrated modeling approach, (b) produced by the spatial index-based approach.

7.5 Application of spatial index-based approach in the upper Ping River basin

Of particular interest was the applicability of the spatial index-based approach for identifying flash flood hazard areas in ungauged basins. This spatial index-based approach consisted of two prime procedures. The first procedure was derivation of FFPI. These FFPI were derived from physiographic variables of a given watershed by using the PCA method. In the upper Nan River, it was found that flow length (FL), distance from stream (Rbuffer), elevation, and drainage density (Dd) were significant variables that contributed to the FFPI. Therefore, it was interesting to know whether these variables were significant factors that can be used to identify areas prone to flash floods in the upper Ping River basin. The second procedure was flash flood hazard mapping. To estimate the degree of flash flood hazards, FFPI values were classified by four classifiers and their corresponding areas were assigned by five degrees of flash flood hazard zones. Thus, it was necessary to examine whether the suitable classifiers that were used to identify flash flood hazard areas for the upper Nan River basin can be used for the upper Ping River basin.

7.5.1 Principal component analysis of physiographic variables

Principal component analysis was used to analyze physiographic variables of the four watersheds located in the upper Ping River basin. The results, as shown in Table 7.12, indicate that FL, Rbuffer, elevation, and Dd were significant variables that strongly contributed to FFPI for identifying flash flood hazard of the Nam Mae Hao watershed (P1), for example. As can be seen from the table, FL, Rbuffer, and elevation are strongly correlated with PC1 and Dd is strongly correlated with PC2. PC1, as mentioned several times, captured the largest variance of the data and PC2 represented the second-largest variance. In other words, the highest weight was given to PC1 and the second-highest weight was given to PC2 for computing FFPI. For the other watersheds, it has been also found that FL, Rbuffer, elevation, and Dd were significant variables contributing to FFPI. This findings are compatible with the analysis based on physiographic variables of the five watersheds located in the upper Nan River basin.

Table 7.12: All principal components and their variance for physiographic variables of the Nam Mae Hao watershed (P1)

Variables	PC1	PC2	PC3	PC4	PC5	PC6
Drainage density	0.11	**0.95**	-0.24	-0.05	-0.13	0.03
Rbuffer	**0.64**	-0.17	-0.17	-0.02	-0.20	**0.71**
Surface slope	0.31	0.17	**0.91**	-0.22	-0.01	-0.03
Elevation	**0.50**	0.02	-0.16	0.05	**0.81**	-0.25
TWI	0.05	0.09	0.20	**0.97**	-0.05	0.03
Flow length	0.48	-0.15	-0.17	0.02	**-0.53**	**-0.66**
Variance [%]	65.03	16.93	8.97	4.77	2.88	1.42
Cumulative variance [%]	65.03	81.96	90.93	95.70	98.58	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

The overall results of PCA based on physiographic variables of the four watersheds located in the upper Ping River basin are summarized in Table 7.13. The physiographic variable having the strongest correlation to PC1 for each watershed is different. This may be consistent with the fact that flash floods are localized phenomenon. Their occurrences are controlled by physiographic features of an individual watershed. However, the results of PCA suggest that FL, Rbuffer, elevation, and Dd were significant variables contributing to FFPI of the four watersheds, as mentioned above.

7.5.2 Comparison of spatial index-based results with satellite images

The spatial index-based approach and its concept for flash flood potential assessment were developed in the upper Nan River basin. Their results were compared and evaluated with results obtained from integrated mathematical models (e.g., hydrologic and hydrodynamic models). On the other hand, none of mathematical models was applied to simulate flash floods in the upper Ping River basins and to identify flash flood hazard areas in its sub-basins. Therefore, simulated flood depths and velocities that are commonly used as

Table 7.13: Summary of mainly contributing variables for the first four PCs of the four study watersheds located in the upper Ping River basin

Watershed	Mainly contributing variables for			
	PC1	PC2	PC3	PC4
Nam Mae Hao (P1)	Rbuffer	Drainage density	Surface slope	TWI
Nam Mae Ngat (P2)	Flow length	Rbuffer	Drainage density	Surface slope
Nam Mae Ngat branch (P3)	Elevation	Drainage density	Rbuffer	Surface slope
Nam Mae Khot (P4)	Rbuffer	Drainage density	Surface slope	TWI

Note: Rbuffer (distance from stream) and TWI (topographic wetness index)

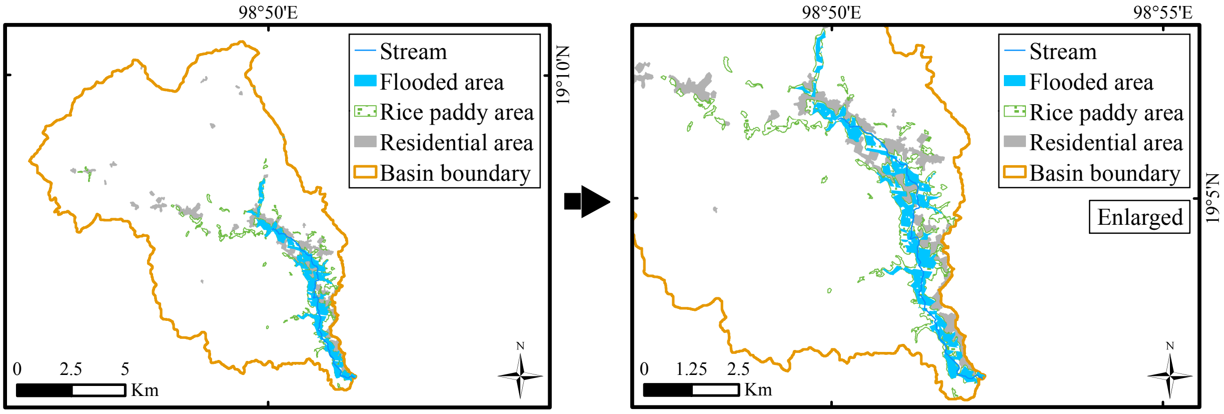
flood hazard indicators were not obtained. In addition, it may be costly to collect data on a high-density spatial distribution of historic flood locations in areas of a few hundred square kilometers through ground observations. Owing to the lack of flood hazard indicators, remote sensing data acquired from the Landsat-7 satellite were used and compared with results of the spatial index-based approach. These remote sensing data can be accessed via the USGS Earth Explorer site (https://earthexplorer.usgs.gov/).

Landsat-7 imagery was processed through the remote sensing technique and was classified into wet and dry areas (described in Section 5.5.2). According to the analysis of the Landsat-7 imagery, it was important to note that clouds and their shadows can lessen the accuracy of inundated area extraction. Moreover, the Landsat-7 satellite observes any given location with a repeat cycle of 16 days that can limit the number of flood events for comparison of inundated areas. From these reasons, two flood events observed by the Landsat-7 satellite on 5 August 2003 and 26 July 2011 were used and compared with results from the spatial index-based approach applied to the four study watersheds of the upper Ping River basin.

Figure 7.25a displays flooded areas extracted from Landsat-7 imagery observed over the Nam Mae Hao watershed (P1) on 5 August 2003 and Figure 7.25b displays a closer view of these extracted flooded areas. Moreover, Figure 7.25c displays flooded areas extracted from Landsat-7 imagery observed over the watershed on on 26 July 2011 and Figure 7.25d displays a closer view of these extracted flooded areas. As can be seen from the figures, extracted flooded areas extended over some parts of rice paddy and residential areas. In addition, it was important to notice that Landsat-7 imagery may capture some parts of rice paddy areas in the watershed between June and September as wet areas because this period is the annual rice cropping season meaning that water may be stored in these areas.

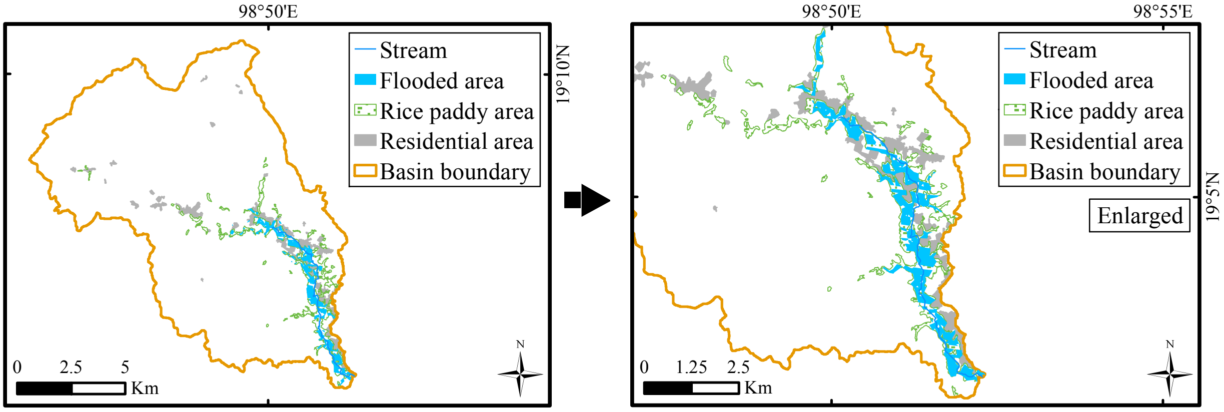
To test the regional classifiers obtained from the analysis of the five watersheds in the upper Nan River basin, a representative probability distribution needed to be determined that well fitted to a FFPI distribution of a given watershed. Interestingly, the GEV probability distribution was best fitted to the FFPI distributions of the four watersheds in the upper Ping River basin. This confirms the findings from the five watersheds that the GEV probability distribution was the representative probability distribution for the combined FFPI distribution of the five watersheds.

For example, the cumulative distribution curve of FFPI values based on physiographic variables of the Nam Mae Hao watershed (P1) are plotted with three commonly used distributions – namely, GEV, Gumbel, and lognormal distributions (Figure 7.26a). It is apparent that the GEV probability distribution is the best fit to the FFPI distribution of the watershed.



(a) Observed on 5 August 2003

(b) Observed on 5 August 2003 (with a closer view)



(c) Observed on 26 July 2011

(d) Observed on 26 July 2011 (with a closer view)

Figure 7.25: Inundated areas extracted from Landsat-7 in the Nam Mae Hao watershed (P1): (a) extracted flooded areas observed on 5 August 2003, (b) extracted flooded areas observed on 5 August 2003 with a closer view, (c) extracted flooded areas observed on 26 July 2011, and (d) extracted flooded areas observed on 26 July 2011 with a closer view.

According to the analysis of the five watersheds in the upper Nan River basin, the 98th percentile of the GEV probability distribution fitted to their FFPI distribution was found to be a “regional” lowest classifier threshold of the watersheds. This regional lowest classifier threshold, as previously mentioned, was used to delineate low to extreme hazard areas of the five watersheds. To examine an applicability of the regional lowest classifier threshold, the 98th percentile of the GEV probability distribution fitted to the FFPI distribution of a given watershed was assigned as a lowest classifier threshold. For example, the probability distribution of the FFPI values based on the Nam Mae Hao watershed (P1) is presented in a histogram as shown in Figure 7.26b. This histogram is compared with the GEV probability density function. The 98th percentile of the GEV probability distribution was used to delineate low to extreme hazard areas of the watershed. These low to extreme hazard areas are displayed in Figure 7.27a. This figure shows that the regional lowest classifier threshold was not suitable to be used as a classifier for the Nam Mae Hao watershed (P1) because the delineated low to extreme hazard areas were much smaller than flooded areas extracted from the Landsat-7 imagery.

Owing to the underestimated hazard areas when the regional lowest classifier threshold was used to delineate low to extreme hazard areas in the Nam Mae Hao watershed (P1), it was necessary to explore a

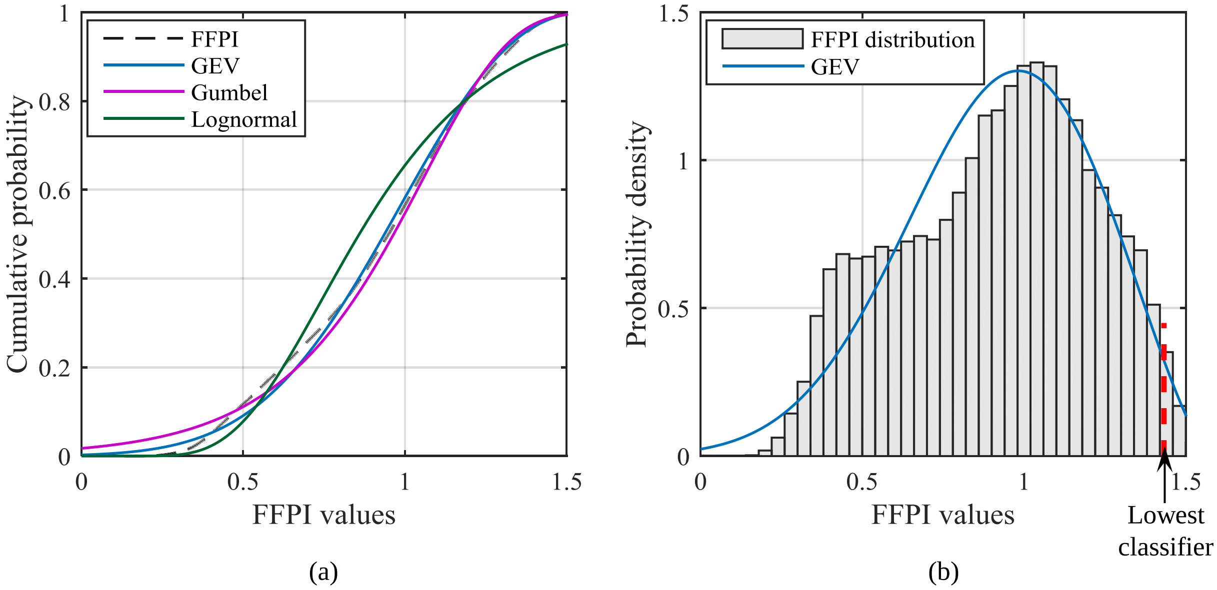
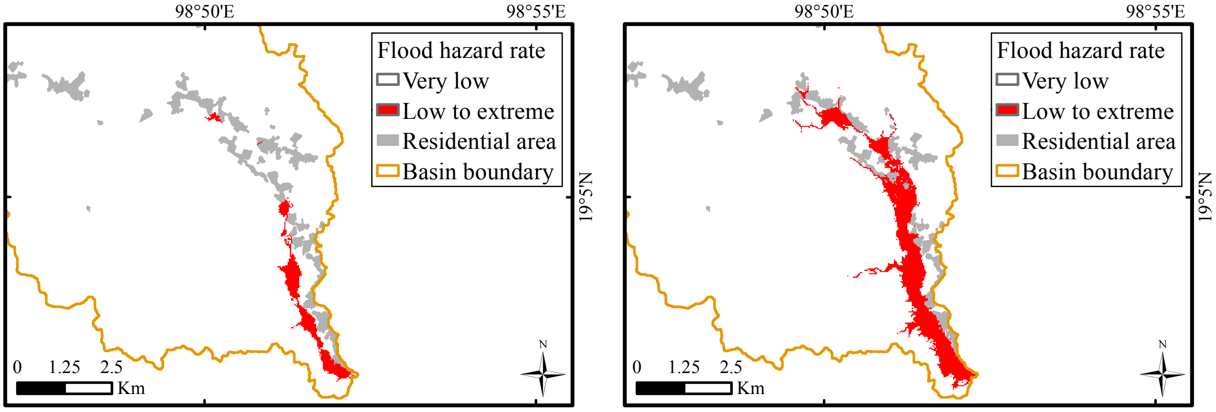


Figure 7.26: FFPI distribution of the Nam Mae Hao watershed (P1) based on PCA and theoretical distribution candidates: (a) cumulative density curves of FFPI and theoretical distribution candidates and (b) generalized extreme value (GEV) probability density function fitted to the distribution of the FFPI values.

lowest classifier threshold for the watershed by comparing these areas with extracted flooded areas based on Landsat-7 observations on 5 August 2003 and 26 July 2011. It was found that the 95th percentile of the GEV probability distribution fitted to the FFPI distribution of the watershed was suitable to be the lowest classifier threshold. Therefore, this lowest classifier threshold was used to delineate low to extreme hazard areas in the Nam Mae Hao watershed (P1) as shown in Figure 7.27b. These areas were compared to the extracted flooded areas of the two flood events as displayed in Figure 7.25.



(a) Lowest classifier threshold (98th percentile)

(b) Lowest classifier threshold (95th percentile)

Figure 7.27: Flash flood hazard maps of the Nam Mae Hao watershed (P1) based on the spatial index-based approach using lowest classifier thresholds: (a) 98th percentile and (b) 95th percentile of the GEV probability distribution.

For the other three watersheds: the Nam Mae Ngat (P2), Nam Mae Ngat branch (P3), and Nam Mae Khot (P4) watersheds, the 95th percentile of the GEV probability distribution was considered suitable to be the lowest classifier threshold. The results of flash flood hazard areas compared with flooded areas extracted from the Landsat-7 imagery for these study watersheds of the upper Ping River basin are presented in Figure A.12 - A.14 (in Appendix A.3.3.2).

In summary, low to extreme hazard areas in the four watersheds of the upper Ping River basin delineated by the regional classifier threshold suitable for the upper Nan River basin were smaller than extracted flooded areas of the two flood events observed by the Landsat-7 satellite. The results indicated that classifiers developed from one region may not be directly used in another. This may be because different characteristics in basins may result in different characteristics of flash floods. Moreover, the size of the flood extent may depend upon the hydraulic and geometric characteristics between a specific river channel and its floodplains. If basins have difference in these characteristics, flood extents may be different. Therefore, classifiers for identifying flash flood hazard areas may be specific to a particular basin.

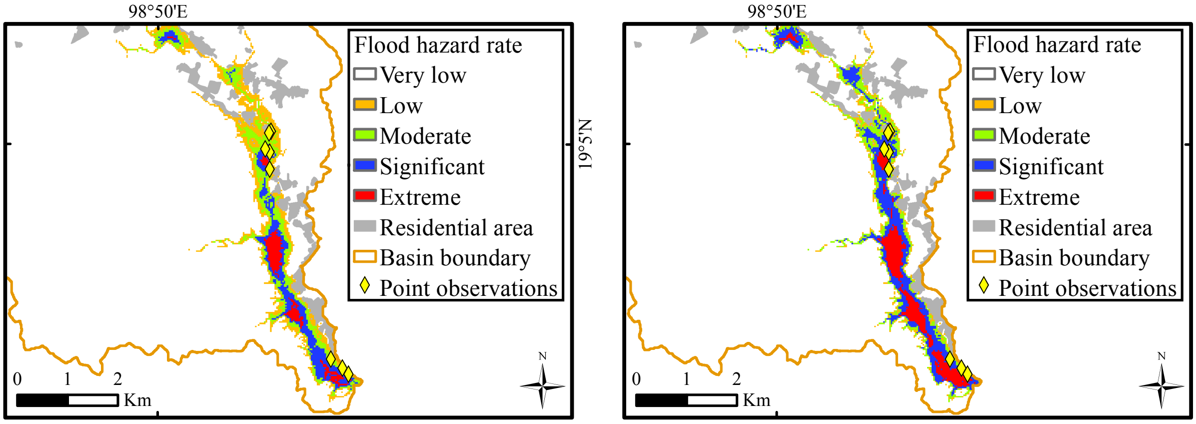
7.5.3 Flash flood hazard maps

As mentioned above in Section 7.5.2, classifiers developed from one region may not be directly used in another. To overcome this problem, it was necessary to adopt the classification procedures for identifying degrees of flash flood hazard. Two commonly used classification methods such as natural breaks and geometrical interval (Dent et al., 2009) were applied to divide the low to extreme hazard areas into four degrees of flash flood hazard: low, moderate, significant, and extreme hazards.

According to the analysis of FFPI values based on the Nam Mae Hao watershed (P1), a suitable lowest classifier threshold for delineating low to extreme hazard areas in the watershed was the 95th percentile of the GEV probability distribution. To divide these low to extreme hazard areas, FFPI values greater than the 95th percentile of the GEV probability distribution were categorized into the four degrees of flash flood hazard by using the natural breaks method. Their corresponding areas were labeled by the degrees of flash flood hazard as shown in Figure 7.28a. This figure shows that point observations of floods (mainly the maximum flood level) collected during the field investigations by the author are coincident to significant, moderate and low flood hazard areas. Interestingly, point observations where maximum flood depths were higher than 0.5 m are identical to significant flood hazard areas.

Furthermore, Figure 7.28b displays a flash flood hazard map of the watershed based on the classification of the FFPI values by using the geometrical interval method. As can be seen from the figure, the point observations are coincident to significant and moderate flood hazard areas. It was found that the significant flood hazard areas cover the point observations where maximum flood depths varied from 0.2 m to 0.5 m. Moreover, some of the point observations where maximum flood depths were lower than 0.5 m were also found in the moderate flood hazard areas. These results suggested that the natural breaks method may be suitable for classifying the FFPI values of the Nam Mae Hao watershed (P1).

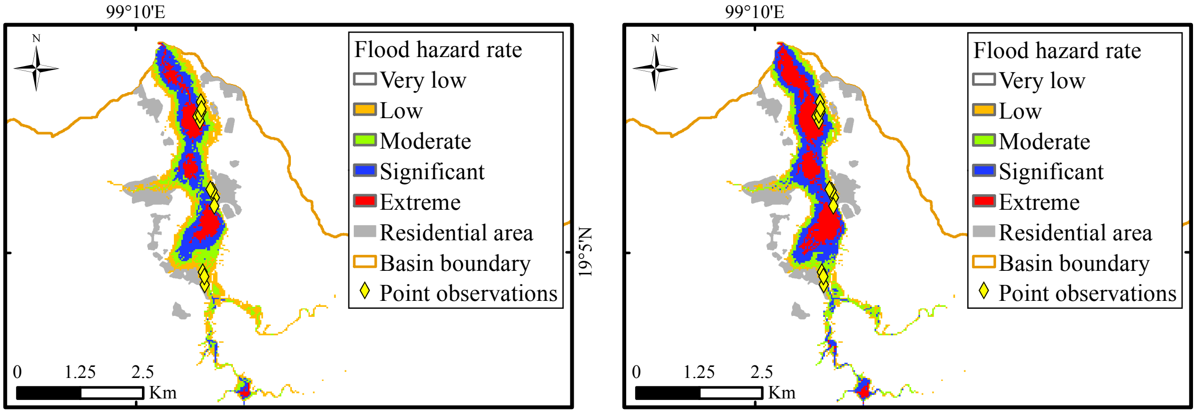
Figure 7.29 gives another example of flash flood hazard maps of the Nam Mae Khot watershed (P4) produced by using the spatial index-based method combined with the two classification methods: natural breaks and geometrical interval classification methods. Flood hazard areas of the Nam Mae Khot watershed (P4) delineated by the classification of FFPI values by using the natural breaks method were compared with point observations of the maximum flood level collected from the field investigations as shown in Figure 7.29a. From this figure it can be seen that the point observations spread over low to extreme hazard areas of the watershed. The point observations where maximum flood depths were higher than 1.0 m were



(a) Natural breaks classification (b) Geometrical interval classification

Figure 7.28: Flash flood hazard maps of the Nam Mae Hao watershed (P1).

found in significant and extreme flood hazard areas. Another point observations where maximum flood depths were between 0.5 m and 1.0 m were found in moderate flood hazard areas. The rest of the point observations where maximum flood depths were lower than 0.5 m were found in low flood hazard areas.



(a) Natural breaks classification (b) Geometrical interval classification

Figure 7.29: Flash flood hazard maps of the Nam Mae Khot watershed (P4).

Moreover, Figure 7.29b illustrates flood hazard areas of the watershed based on the geometrical interval classification method and the point observations of the maximum flood level. This figure shows that the point observations are compatible with extreme, significant and moderate flood hazard areas. The point observations where maximum flood depths were higher than 1.0 m were found in significant and extreme flood hazard areas. This was consistent with the results based on the natural breaks classification method. However, a difference was that some of the point observations where maximum flood depths were between 0.5 m and 1.0 m were found in significant flood hazard areas delineated by the geometrical interval classification method.

Taken together, the results above suggested that the spatial index-based method combined with the natural breaks classification method can be a useful tool for delineating flash flood hazard areas where flood observations are limited.

7.6 Summary of results

Synthetic rainfall series were obtained from the combination of the temporal rainfall pattern and regional rainfall frequency relationships. From this, a temporal rainfall pattern of the upper Nan River was derived by averaging 13 rainstorms of different rainfall durations in rainy months of 2008 and 2011. The temporal rainfall pattern of the study region was represented by a dimensionless mass curve that shows relationships of dimensionless cumulative rainfall depth and percentage of total rainstorm duration. Then, the regional rainfall frequency relationships were derived. These rainfall frequency relationships were established by using the Generalized Pareto (GPA) distribution that was found to be the most suitable probability distribution for the study region according to the *L*-moment method.

The PANTA RHEI hydrologic model was employed to characterize flash floods in five sub-basins of the upper Nan River basin – namely, Nam Rim watershed (N1), Nam Hui watershed (N2), Nam Lae watershed (N3), Nam Pua watershed (N4), and Nam Khwang watershed (N5). Owing to the lack of reliable runoff data in the five sub-basins of the upper Nan River basin, it was necessary to analyze historic extreme flood events, which occurred in the upper Nan River basin. It was found that the performance of PANTA RHEI model was strongly improved when sub-daily rainfall time series were used as an input. As a result, the PANTA RHEI model was capable to reasonably well reproduce the extreme flood events of 2008 and 2011 based on the performance criteria such as efficiency index and relative peak error. To simulate flash floods in the five sub-basins, the calibrated model parameters of the upper Nan River basin were transferred to these five sub-basins. It is important to note that the transfer of calibrated model parameters from the parent basin to its sub-basins can be accepted if the model parameters are not too sensitive. Therefore, six significant model parameters from three main components of the runoff generation in PANTA RHEI were selected and examined using a sensitivity analysis with respect to simulated variability. Three of the parameters were maximum interception storage, a factor of potential evapotranspiration, and a factor of percolation in the runoff formation component. Two of the parameters were a factor of interflow storage and a factor of groundwater storage in the runoff concentration component. The last parameter that was a flow velocity factor was selected from the runoff routing component. According to the sensitivity analysis of the model parameters, a small difference of about ±7% was found between simulated peak floods based on the calibrated model parameters and simulated peak floods based on values of the model parameters, which varied within ±20% of the values of the calibrated model parameters. Transferability of the model parameters was therefore considered to be acceptable.

The MEADFLOW hydrodynamic model was applied to estimate flash flood extents, depths and velocities in the downstream areas of the five sub-basins. The MEADFLOW model was calibrated by comparing its estimated flood depths with high-flood marks, which occurred during the 2011 flood event. The overall results were acceptable on the basis of statistical measures such as efficiency index and *RMSE*.

The spatial index-based approach delivered flash flood potential indices (FFPI) by using principal component analysis (PCA). In this research work, six physiographic variables – namely, drainage density, distance from stream, surface slope, elevation, topographic wetness index, and flow length, were identified to be relevant to derive the FFPI. On the basis of PCA, flow length, distance from stream, elevation, and drainage density were significant variables that contributed to the FFPI.

For sensitivity analysis of the spatial index-based approach, effects of different spatial rainfall distributions based on the data of the Nam Hui watershed (N2) were examined. Results of this sensitivity analysis shown a variation of FFPI distributions in cases of spatial rainfall distributions computed by using raingauge data interpolation. In some of the cases, higher values of the FFPI were delineated in upstream areas of the watershed, whereas lower values were assigned to downstream areas. These cases pointed out a drawback of the PCA method because it cannot to transform the rainfall to dynamic runoff processes like the integrated modeling approach. Another case was that the highest value of rainfall was given to a raingauge near the middle part of the watershed. As a result, higher FFPI values were found in the middle part of the watershed. It can be seen that the spatial index-based approach was significantly sensitive to spatial rainfall distributions. What is more, some spatial rainfall distributions interpolated by using the Thiessen polygon method can disturb the variability of data pattern that may lead to a mistake interpretation of PCA. Thus, it was concluded that dynamic variables such as rainfall cannot be combined with or analyzed together with static variables such as the physiographic variables by using PCA.

However, an extension of the “static” FFPI value of an endangered zone towards a “dynamic” value by combining it with an actual quantitative forecast of extreme meteorological conditions can be a field of future research. In doing so, the available “static” FFPI may be a valuable indicator to become a representative indicator of flash flood occurrence in a (real-time) forecast.

According to classification of FFPI, the FFPI distribution of the five sub-basins was best fitted to the generalized extreme value (GEV) distribution. On the basis of the five sub-basins, four classifiers were assigned in order to group the FFPI values and identify areas prone to flash floods into five levels of flash flood hazard: very low, low, moderate, significant, and extreme. The lowest classifier was used to determine a flooding boundary. Areas outside of the boundary had FFPI values lower than the lowest classifier. These areas were assigned as a very low flood hazard because the occurrence of flash floods is extremely rare on the basis of the field investigations. In contrast, areas inside of the boundary were compared with flood extent simulated by using the integrated modeling approach. It was found that the 98th percentile of the GEV probability distribution fitted to the combined FFPI distributions based on the five sub-basins was suitable to be the lowest classifier of the upper Nan River basin. Here, the 98th percentile of the GEV probability distribution is called the “regional” lowest classifier. In order to avoid confusing use of this regional lowest classifier, the percentile does not refer to an annual nonexceedance probability for a series of events (i.e., the 98th percentile means a flood with a return period of 50 years). However, the percentile refers to the position of an individual on a given probability density distribution.

For further classification of FFPI based on the five sub-basin of the upper Nan River basin, the areas having FFPI values higher than the regional lowest classifier were categorized by the rest of the classifiers into low, moderate, significant, and extreme hazard areas. These four hazard areas were compared with flash flood hazard maps under different rainfall scenarios produced by using the integrated modeling approach. For example, the 98.30th, 98.45th, and 99.57th percentiles of the GEV probability distribution fitted to the FFPI distribution of the Nam Rim watershed were found to provide a fair agreement of flash flood hazard maps of the R130D6 rainfall scenario. The R130D6 scenario means an event where a rainfall amount of 130 mm was assumed to be spatially uniformly distributed over a given watershed for a duration of six hours. Another set of classifiers that was the 98.25th, 98.40th, and 99.45th percentiles of the GEV probability distribution was found to be suitable to classify FFPI values for identifying flash flood hazard areas that are consistent with ones of the R200D6 rainfall scenario produced by the modeling approach. The findings suggest that this classification based on an identified index distribution is useful because its classifiers are adjustable. This could potentially lead to flash flood potential assessment of varying rainfall conditions.

The spatial index-based approach can be applied everywhere but the quality of results depends upon validation conditions. If it is an existing condition that a mathematical model (i.e., a combined hydrologic and hydrodynamic model) can provide highly detailed flash flood hazard areas (i.e., low, moderate, significant, and extreme hazards), results of the spatial index-based approach can be validated with these areas. Moreover, the spatial index-based approach can produce flash flood hazard maps under different rainfall scenarios by adjusting classifiers according to flash flood hazard maps produced by the mathematical model. On the other hand, conditions of ungauged watersheds may limit effective use of the mathematical model in flash flood hazard assessment. Therefore, highly detailed flash flood hazard areas usually do not exist in ungauged watersheds to compare with results obtained by the spatial index-based approach. Nonetheless, The spatial index-based approach was applied to four ungauged watersheds located in the upper Ping River basin. To examine transferability of classifiers found in the analysis of the upper Nan River basin, the regional lowest classifier was used to delineate flood boundaries in the four watersheds of the upper Ping River basin. The areas inside of the boundaries were smaller than the flood extent extracted from the Landsat-7 imagery. These findings indicated that classifiers developed from one region may not be directly used in another. This may be due to different characteristics in basins, which may result in different characteristics of flash floods. In addition, the size of the flood extent may depend upon the hydraulic and geometric characteristics between a specific river channel and its floodplains. If basins have difference in these characteristics, flood extents may be different. Therefore, classifiers for identifying flash flood hazard areas may be specific to a particular basin. However, the spatial index-based approach has proved to a robust tool for multivariate analysis and weight in determining FFPI because it analyzes and weights variables according to their influences on occurrence of flash floods without involving "expert" opinion.

Although there are some limitations of the spatial index-based approach, an application of commonly used classification methods such as natural breaks and geometrical intervals methods could be used to classify FFPI values inside of the boundaries into low, moderate, significant, and extreme hazards. The results based on the field investigations of this dissertation suggest that the spatial index-based approach with the help of the natural breaks classification method can be a useful tool for mapping flash flood hazard areas where flood observations are limited.

8 Discussion, Conclusions, and Outlooks

8.1 Discussion

An essential objective of this research work was to contribute to a better understanding of flash flood potential assessment under poor data availability. The research work was focused on headwater basins with drainage areas less than about 350 km^2 because these basins, in general, have a high potential for flash flood occurrence (Marchi et al., 2010). At this basin scale, observed runoff data are often not available, especially in developing and emerging countries (e.g., Thailand and Vietnam). Besides the lack of runoff measurements, flash flood potential assessment in these countries is even more difficult because reliable rainfall data at high temporal resolutions (e.g., hourly) and spatial resolutions (e.g., a few km^2) are usually missing. In some regions, high temporal rainfall data are available but available periods are generally too short for flood planning and management (Gaume et al., 2007). Under data scarcity conditions, a complex mathematical model for the simulation of flash floods (i.e., a combined hydrologic and hydrodynamic model) is difficult to apply. Its results can be highly uncertain (Di Baldassarre, 2012). To assess flash flood potential under such conditions, interdisciplinary approaches need to be used for comparison purposes.

The author's attention was concentrated not only on scientific issues but also on developing practical tools for flash flood risk assessment and management. For the public exposed to flash flood risk, spatial visualization and information of flash flood hazard assessment (i.e., a flash flood hazard map) is more useful. As mentioned above, challenges of flash flood hazard assessment are due to the lack of reliable runoff data; results from different approaches are necessary to be compared and discussed. In this dissertation, an integrated hydrologic and hydrodynamic modeling approach and a spatial index-based approach were used to spatially assess flash flood hazard in ungauged sub-basins in tropical regions, in particular in Thailand.

Firstly, a framework of integrated hydrologic and hydrodynamic models was elaborated to predict well the flood extent, even though the data on flash flooding were scare in the study sub-basins. A highlight advantage of the framework is that useful information from a larger scale can be transmitted to simulate flash floods at a smaller scale. One disadvantage is the high input data demand. In the framework, a distributed hydrologic model (i.e., PANTA RHEI) was used to reproduce extreme river floods in the parent catchment of the ungauged sub-basins. At a larger scale (i.e., a few thousand km^2), observed river flow data are usually available. Therefore, PANTA RHEI was calibrated and validated against the river flow data recorded during two extreme flood events of 2011 and 2008, respectively, and was proved its strong capability to reproduce these events. To simulate flash floods in the ungauged sub-basins, the calibrated model parameters of their parent catchment were transferred to these sub-basins. Transferability of the model parameters was proved to be acceptable by using a sensitivity analysis of the model parameters with respect to simulated variability. In addition, the parent catchment and its sub-basins were considered to be homogeneous in an extreme climate condition based on regional frequency analysis of annual maximum daily rainfall data. This confirms the transferability of the model parameters under the underlying assumption of regionalization concepts that basins similar in climatology and physiography will behave similarly from a hydrologic point of view (Wagener et al., 2013). In the other part of the framework, results from the hydrologic model were used to drive a hydrodynamic model (i.e., MEADFLOW). The MEADFLOW model was applied to spatially route runoff simulated from the hydrologic model in downstream areas of the sub-basins. Owing to the lack of runoff and stage measurements, downstream boundary conditions of the MEADFLOW model were assigned by the bankfull elevation. The MEADFLOW model was calibrated by adjusting roughness coefficients and comparing its simulated flood depths with high-flood marks, which occurred during the 2011 flood event. The overall results were acceptable. However, the simulations of flash floods in the sub-basins could not be completely validated because of two main reasons: (i) the lack of runoff data in these sub-basins and (ii) a limited number of point observations of flash floods collected during the field investigations by the author. Therefore, uncertainty must be taken into account in the interpretation of results.

Secondly, a new spatial index-based approach was developed. It has proved to be very useful for flash flood hazard assessment in ungauged basins. This spatial index-based approach delivers flash flood potential indices (FFPI) by using principal component analysis (PCA). On the basis of PCA, physiographic variables that have potential to flash flood occurrence are flow length, distance from stream, elevation, and drainage density. These findings are consistent in good agreement with other studies, which have shown that the accumulated flow (sometimes referred to as flow length) and elevation variables are significant factors in indicating a flash flood potential (Kazakis et al., 2015; Nandi et al., 2016). The reason for this is that high flow contributing areas and lowlands are most likely to be flooded. Moreover, the distance from stream and slope variables are considered as significantly contributing factors. These substantiate previous findings in the literature (e.g., Kazakis et al., 2015). The highest number of major floods occurred in areas, which are in close proximity to streams and in areas with almost flat to mild slopes (Nandi et al., 2016). High correlations between drainage density and the FFPI are found in this dissertation. This finding is consistent with Merz and Blöschl (2005), who stated that drainage density is correlated to flood peaks.

In some recent studies (e.g., Kazakis et al., 2015; Khosravi et al., 2016), index-based approaches were extended by including rainfall variables into statistical analysis to construct flash flood indices. Even though these studies presented success in identifying areas prone to flash floods, their analyzes seem to have some weaknesses. In Kazakis et al. (2015)'s study, flood hazard indices were computed by using an analytic hierarchy process (AHP), which involves "expert" knowledge. In Khosravi et al. (2016)'s work, weighting factors for physiographic and rainfall variables were calculated by statistical methods with comparison to historic flood locations. Because statistical methods do not account for overland flow routing, the rainfall at a given grid was weighted with respect to flooding or non-flooding observation points. It happened that some higher amounts of rainfall were assigned as a weight of zero. This is because there was no flood observation point or no flood occurrence on the same grids as the higher amounts of rainfall detected. Therefore, these higher amounts of rainfall were considered by the statistical methods as insignificant factors to cause the occurrence of flash floods.

Because of the uncertainties and partially criticizable procedures used in the above-mentioned studies, the author investigated effects of different spatial rainfall distributions on the developed index-based approach. Unlike other studies carried out in this research area, results of the developed index-based approach suggest that rainfall variables should not be analyzed together with other physiographic variables by using statistical methods such as PCA. This is due to the straightforward procedure of PCA that does not involve "expert" judgment. PCA spatially analyzes variables and automatically weights them according to their influences on the occurrence of flash floods. Moreover, statistical methods (e.g., PCA) cannot take into account the rainfall-runoff transformation and runoff routing. For example, if higher amounts of rainfall spatially distribute over upstream areas of a given basin, results of PCA may show that these upstream areas have higher potentials of flash flood occurrence but flash floods often occur in downstream areas as their quick movements from higher to lower areas.

Although the developed index-based approach at this present stage is not include rainfall variables, it can provide reasonable results on flash flood potential assessment. In a future study, rainfall variables will be included into the developed index-based approach again but in a different way.

8.2 Conclusions

The main objective of this dissertation is to improve the understanding of flash flood potential assessment under poor data availability. This poor data situation usually limits research on flash floods, especially in developing and emerging countries. Thus, intensive guidelines for assessing flash flood potential and mapping flash flood hazard are needed. To assess flash flood potential in ungauged basins, two approaches were developed and used to produce flash flood hazard maps for comparison purposes.

The first approach is an integrated hydrologic and hydrodynamic modeling approach. The hydrologic model (i.e., PANTA RHEI) was applied to one of the two selected watersheds of ungauged basins. The capability of PANTA RHEI for reproducing extreme flood events reasonably performs when hourly rainfall series are input. In addition, results of the hydrologic model are supported by the underlying assumption of parameter regionalization that basins similar in climatology and physiography will behave similarly from a hydrologic perspective. The study basins passed a regional homogeneity test and were considered to be homogeneous in an extreme climate condition. *L*-moment analysis confirmed that the Generalized Pareto (GPA) distribution provides good fit to the observed data of annual maximum daily rainfall. This *L*-moment analysis is considered to be a practical and robust tool for identifying the most suitable probability distribution of a region. In addition, sensitivity analysis of PANTA RHEI model parameters with respect to simulated variability seems to be necessary for flash flood modeling in ungauged basins. Results of the PANTA RHEI model were combined with the hydrodynamic model (i.e., MEADFLOW) in order to estimate indicators (i.e., flood extent, depth of flooding, and velocity of floodwater). These useful indicators are usually used to estimate degrees of flash flood hazard.

The second approach is a new spatial index-based approach developed by using the results of the model-based procedure as a reference. This approach delivers flash flood potential indices (FFPI) by using the principal component analysis (PCA). PCA spatially analyzes physiographic variables – namely, drainage density, distance form stream, surface slope, elevation, topographic wetness index, and flow length and combines their information to construct FFPI. The main strengths of this spatial index-based approach are as follows: (i) the FFPI is derived from available data, (ii) cost-effectiveness of the application, and (iii) objectivity because of the automatically weighting technique (i.e., PCA). The results of this research work can be a useful aid for decision makers because the spatial index-based approach provides a user-friendly visualized information of flood hazard potential. Moreover, the classification of FFPI values shows that the spatial index-based approach has led to a greater degree of flexibility and applicability in flash flood hazard mapping. By changing classifiers, areas of a given flash flood hazard can be changed according to various exceedance probabilities of events. In total, the spatial index-based approach is considered to be an effective procedure for flash flood hazard assessment. In addition, the spatial index-based approach can be extended to cope with flash flood occurrence in a (real-time) forecast if routing the effective rainfall is integrated with the approach.

The findings of this dissertation provide insights for the use of statistical methods in hydrology. Statistical methods are attractive because they can use functions or statistical relationships to fit with observed data without taking into account the hydrologic response. Therefore, the interpretation of results obtained by the statistical methods should be also interpreted under consideration of hydrologic credibility.

8.3 Outlooks

In this dissertation, the effects of the "static" physiographic characteristics for flash flood potential assessment in ungauged basins had been extensively studied. For this, the new spatial index-based approach is introduced to compute flash flood potential indices (FFPI) from the static physiographic characteristics, which are mainly not changing with time. At the present stage, the new spatial index-based approach does not include non-static parameters like observed on-line rainfall or forecasted rainfall. However, this is the first step to go into further research. Therefore, the author suggests the following interesting research directions for improving flash flood potential assessment:

- According to the spatial index-based approach, the percentile of the GEV distribution (fitted to the distribution of FFPI values of a given sub-basin) for flash flood boundary identification slightly varies from one sub-basin to another sub-basin. Thus, research should be done in future to provide a method of fine tuning of the variation of the above-mentioned percentile. In doing so, more case studies and data collection should be conducted.

- This research work can be extended towards a (real-time) forecast. For real-time forecasting of flash flood potential, non-static variables such as actual and forecasted rainfall intensity and soil moisture conditions should be included. For example, an analytical routing method as a function of non-static variables and runoff characteristics on grid basis should be developed. This analytical routing method can be used to transform spatial rainfall intensity distributions to variation in runoff depth. This variation can be integrated with results of the "static" FFPI in order to provide real-time forecasting of flash flood potential. Therefore, a method to include these static and non-static variables should be sought for improving the accuracy of potential flash flood forecasting in future research.

- The accuracy of flash flood potential assessment based on the spatial index-based approach may be improved with more satellite-based observations of, for example, inundation areas, soil moisture of the upper soil layer, and vegetation. As satellite-based observations have been improving in both temporal and spatial scales, these observations can be used to validate and update flash flood hazard maps and improve the understanding of flash floods.

References

1. Abdelkareem, M. (2017): *Targeting flash flood potential areas using remotely sensed data and GIS techniques*. Natural Hazards, 85, Nr. 1, 19–37.
2. Abt, S. R., R. J. Wittier, A. Taylor and D. J. Love (1989): *Human stability in a high flood hazard zone*. Water Resources Bulletin, 25, Nr. 4, 881–890.
3. Adamovic, M., F. Branger, I. Braud and S. Kralisch (2016): *Development of a data-driven semi-distributed hydrological model for regional scale catchments prone to Mediterranean flash floods*. Journal of Hydrology, 541, 173–189.
4. Amponsah, W., L. Marchi, D. Zoccatelli, G. Boni, M. Cavalli, F. Comiti, S. Crema, A. Lucía, F. Marra and M. Borga (2016): *Hydrometeorological characterization of a flash flood associated with major geomorphic effects: Assessment of peak discharge uncertainties and analysis of the runoff response*. Journal of Hydrometeorology, 17, Nr. 12, 3063–3077.
5. Anhalt, M. and G. Meon (2008): *Risk-based procedure for design and verification of dam safety*. In S. Simonovic, P. Bourget und F. Blanchard (Ed.): *Proceedings of the 4th International Symposium on Flood Defence.*, 154(1)–154(8).
6. Artigue, G., A. Johannet, V. Borrell and S. Pistre (2012): *Flash flood forecasting in poorly gauged basins using neural networks: Case study of the Gardon de Mialet basin (southern France)*. Natural Hazards and Earth System Science, 12, Nr. 11, 3307–3324.
7. Arvidson, T., J. Gasch and S. N. Goward (2001): *Landsat 7's long-term acquisition plan – an innovative approach to building a global imagery archive*. Remote Sensing of Environment, 78, Nr. 1-2, 13–26.
8. Awadallah, A. G., M. ElGamal, A. ElMostafa and H. ElBadry (2011): *Developing intensity-duration-frequency curves in scarce data region: an approach using regional analysis and satellite data*. Engineering, 3, Nr. 3, 215–226.
9. Bajabaa, S., M. Masoud and N. Al-Amri (2014): *Flash flood hazard mapping based on quantitative hydrology, geomorphology and GIS techniques (case study of Wadi Al Lith, Saudi Arabia)*. Arabian Journal of Geosciences, 7, Nr. 6, 2469–2481.
10. Ballesteros-Cánovas, J. A., M. Eguibar, J. M. Bodoque, A. Díez-Herrero, M. Stoffel and I. Gutiérrez-Pérez (2011): *Estimating flash flood discharge in an ungauged mountain catchment with 2D hydraulic models and dendrogeomorphic palaeostage indicators*. Hydrological Processes, 25, Nr. 6, 970–979.
11. Baumgartner, A. and H. J. Liebscher (1996): *Allgemeine Hydrologie*. Vol. 1, Lehrbuch der Hydrologie. 2. Edition. Berlin: Borntraeger.
12. Bell, F. G. (1999): *Geological Hazards: Their assessment, avoidance, and mitigation*. London: E & FN Spon.

13. BELL, F. (1969): *Generalized rainfall - duration - frequency relationships*. Journal of the Hydraulics Division, 95, Nr. 1, 311–328.
14. BEVEN, K. (2001): *How far can we go in distributed hydrological modelling?* Hydrology and Earth System Sciences, 5, Nr. 1, 1–12.
15. BEVEN, K. (2012): *Rainfall-Runoff Modelling: The Primer*. 2. Edition. New York: John Wiley & Sons Inc.
16. BEVEN, K. J. and M. J. KIRKBY (1979): *A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant*. Hydrological Sciences Bulletin, 24, Nr. 1, 43–69.
17. BIEGER, K., H. RATHJENS, P. M. ALLEN and J. G. ARNOLD (2015): *Development and evaluation of bankfull hydraulic geometry relationships for the physiographic regions of the United States*. Journal of the American Water Resources Association, 51, Nr. 3, 842–858.
18. BLÖSCHL, G. and M. SIVAPALAN (1995): *Scale issues in hydrological modelling: A review*. Hydrological Processes, 9, Nr. 3-4, 251–290.
19. BO, Z., S. ISLAM and E. A. B. ELTAHIR (1994): *Aggregation-disaggregation properties of a stochastic rainfall model*. Water Resources Research, 30, Nr. 12, 3423–3435.
20. BONNIFAIT, L., G. DELRIEU, M. LE LAY, B. BOUDEVILLAIN, A. MASSON, P. BELLEUDY, E. GAUME and G.-M. SAULNIER (2009): *Distributed hydrologic and hydraulic modelling with radar rainfall input: Reconstruction of the 8–9 September 2002 catastrophic flood event in the Gard region, France*. Advances in Water Resources, 32, Nr. 7, 1077–1089.
21. BORGA, M., P. BOSCOLO, F. ZANON and M. SANGATI (2007): *Hydrometeorological analysis of the 29 August 2003 flash flood in the Eastern Italian Alps*. Journal of Hydrometeorology, 8, Nr. 5, 1049–1067.
22. BORGA, M. L. V., E. GAUME, M. MARTINA and J. THIELEN (2008): *Realtime Guidance for Flash Flood Risk Management: FLOODsite report, T16-08-02*. UniPad and HR Wallingford.
23. BRAUD, I., H. ROUX, S. ANQUETIN, M. MAUBOURGUET, C. MANUS, P. VIALLET and D. DARTUS (2010): *The use of distributed hydrological models for the Gard 2002 flash flood event: Analysis of associated hydrological processes*. Journal of Hydrology, 394, Nr. 1-2, 162–181.
24. BRAUER, C. C., A. J. TEULING, A. OVEREEM, VAN DER VELDE, Y., P. HAZENBERG, WARMERDAM, P. M. M. and R. UIJLENHOET (2011): *Anatomy of extraordinary rainfall and flash flood in a Dutch lowland catchment*. Hydrology and Earth System Sciences, 15, Nr. 6, 1991–2005.
25. BRYNDAL, T. (2014): *A method for identification of small Carpathian catchments more prone to flash flood generation, based on the example of south-eastern part of the Polish Carpathians*. Carpathian Journal of Earth and Environmental Sciences, 9, Nr. 3, 109–122.
26. CARPENTER, T. M., J. A. SPERFSLAGE, K. P. GEORGAKAKOS, T. SWEENEY and D. L. FREAD (1999): *National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems*. Journal of Hydrology, 224, Nr. 1-2, 21–44.
27. CASTRO, J. M. and P. L. JACKSON (2001): *Bankfull discharge recurrence intervals and regional hydraulic geometry relationships: Patterns in the Pacific Northwest, USA*. Journal of the American Water Resources Association, 37, Nr. 5, 1249–1262.

28. Chow, V. T. (1959): *Open-Channel Hydraulics*. Boston: McGraw-Hill.

29. Chow, V. T., D. R. Maidment and L. W. Mays (1988): *Applied Hydrology*. New York: McGraw-Hill.

30. Congalton, R. G. (1991): *A review of assessing the accuracy of classifications of remotely sensed data*. Remote Sensing of Environment, 37, Nr. 1, 35–46.

31. Connolly, R. D., J. Schirmer and P. K. Dunn (1998): *A daily rainfall disaggregation model*. Agricultural and Forest Meteorology, 92, Nr. 2, 105–117.

32. Costa, J. E. and R. D. Jarrett (2008): *An evaluation of selected extraordinary floods in the United States reported by the U.S. Geological Survey and implications for future advancement of flood science*. United States Geological Survey, Scientific Investigations Report 2008-5164.

33. Costache, R. and L. Zaharia (2017): *Flash-flood potential assessment and mapping by integrating the weights-of-evidence and frequency ratio statistical methods in GIS environment – case study: Bâsca Chiojdului River catchment (Romania)*. Journal of Earth System Science, 126, Nr. 4, 1–19.

34. Coulibaly, P., F. Anctil and B. Bobée (2000): *Daily reservoir inflow forecasting using artificial neural networks with stopped training approach*. Journal of Hydrology, 230, Nr. 3-4, 244–257.

35. Cox, R. J., T. D. Shand and M. J. Blacka (2010): *Australian Rainfall and Runoff: Revision Project 10: Appropriate Safety Criteria for People*. Water Research Laboratory, The University of New South Wales.

36. Dalrymple, T. and M. A. Benson (1968): *Measurement of peak discharge by the slope-area method*. United States Geological Survey, Techniques of Water-Resources Investigation of the United States Geological Survey.

37. Dawod, G. M., M. N. Mirza and K. A. Al-Ghamdi (2011): *GIS-based spatial mapping of flash flood hazard in Makkah City, Saudi Arabia*. Journal of Geographic Information System, 3, 225–231.

38. de Almeida, I. K., A. K. Almeida, J. A. A. Anache and T. A. Sobrinho (2014): *Estimation on time of concentration of overland flow in watershed: A review*. Geociencias, 33, Nr. 4, 661–671.

39. Dent, B. D., J. Torguson and T. W. Hodler (2009): *Cartography: Thematic Map Design*. 6. Edition. New York: McGraw-Hill.

40. Di Baldassarre, G. (2012): *Floods in a Changing Climate: Inundation Modelling*. Cambridge: Cambridge University Press, International hydrology series.

41. Diakakis, M. (2011): *A method for flood hazard mapping based on basin morphometry: Application in two catchments in Greece*. Natural Hazards, 56, Nr. 3, 803–814.

42. Domeneghetti, A., S. Vorogushyn, A. Castellarin, B. Merz and A. Brath (2013): *Probabilistic flood hazard mapping: Effects of uncertain boundary conditions*. Hydrology and Earth System Sciences, 17, Nr. 8, 3127–3140.

43. Dooge, J. C. I. (1959): *A general theory of the unit hydrograph*. Journal of Geophysical Research, 64, Nr. 2, 241–256.

44. Dreyfus, G. (2005): *Neural Networks: Methodology and Applications*. Berlin: Springer-Verlag.

45. DUNNE, T. and L. B. LEOPOLD (1978): *Water in Environmental Planning*. New York: W. H. Freeman.

46. DUNTEMAN, G. H. (1989): *Principal Components Analysis*. Newbury Park: Sage Publications.

47. DVWK (1991): *Hydraulische Berechnung von Fliessgewässern: DK 551.51/54 Fliessgewässer ; DK 532.543 Hydraulik*. Vol. 220, DVWK-Merkblätter zur Wasserwirtschaft. Hamburg and Berlin: Paul Parey.

48. EL BASTAWESY, M., K. WHITE and A. NASR (2009): *Integration of remote sensing and GIS for modelling flash floods in Wadi Hudain catchment, Egypt*. Hydrological Processes, 23, Nr. 9, 1359–1368.

49. EL-MAGD, I. A., E. HERMAS and M. EL BASTAWESY (2010): *GIS-modelling of the spatial variability of flash flood hazard in Abu Dabbab catchment, Red Sea Region, Egypt*. The Egyptian Journal of Remote Sensing and Space Science, 13, Nr. 1, 81–88.

50. ESTUPINA-BORRELL, V., D. DARTUS and R. ABABOU (2006): *Flash flood modeling with the MARINE hydrological distributed model*. Hydrology and Earth System Sciences Discussions, 3, Nr. 6, 3397–3438.

51. FABRY, F. and I. ZAWADZKI (1995): *Long-Term Radar Observations of the Melting Layer of Precipitation and Their Interpretation*. Journal of the Atmospheric Sciences, 52, Nr. 7, 838–851.

52. FERNÁNDEZ, D. S. and M. A. LUTZ (2010): *Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis*. Engineering Geology, 111, Nr. 1-4, 90–98.

53. FOODY, G. M., E. M. GHONEIM and N. W. ARNELL (2004): *Predicting locations sensitive to flash flooding in an arid environment*. Journal of Hydrology, 292, Nr. 1-4, 48–58.

54. GALLAGHER, A. S. (1999): *Drainage basins*. In M. B. BAIN und N. J. STEVENSON (Ed.): *Aquatic Habitat Assessment*. Bethesda, Md.: American Fisheries Society, 25–34.

55. GARAMBOIS, P. A., H. ROUX, K. LARNIER, D. LABAT and D. DARTUS (2015): *Parameter regionalization for a process-oriented distributed model dedicated to flash floods*. Journal of Hydrology, 525, 383–399.

56. GAUME, E., V. BAIN, P. BERNARDARA, O. NEWINGER, M. BARBUC, A. BATEMAN, L. BLAŠKOVIČOVÁ, G. BLÖSCHL, M. BORGA, A. DUMITRESCU, I. DALIAKOPOULOS, J. GARCIA, A. IRIMESCU, S. KOHNOVA, A. KOUTROULIS, L. MARCHI, S. MATREATA, V. MEDINA, E. PRECISO, D. SEMPERE-TORRES, G. STANCALIE, J. SZOLGAY, I. TSANIS, D. VELASCO and A. VIGLIONE (2009): *A compilation of data on European flash floods*. Journal of Hydrology, 367, Nr. 1-2, 70–78.

57. GAUME, E. and M. BORGA (2008): *Post-flood field investigations in upland catchments after major flash floods: Proposal of a methodology and illustrations*. Journal of Flood Risk Management, 1, Nr. 4, 175–189.

58. GAUME, E., M. LIVET, M. DESBORDES and J.-P. VILLENEUVE (2004): *Hydrological analysis of the river Aude, France, flash flood on 12 and 13 November 1999*. Journal of Hydrology, 286, Nr. 1-4, 135–154.

59. GAUME, E., N. MOUHOUS and H. ANDRIEU (2007): *Rainfall stochastic disaggregation models: calibration and validation of a multiplicative cascade model*. Advances in Water Resources, 30, Nr. 5, 1301–1319.

60. GELLESZUN, M., P. KREYE and G. MEON (2015): *Lexikografische Kalibrierungsstrategie für eine effiziente Parameterschätzung in hochaufgelösten Niederschlag-Abfluss-Modellen: Hydrologie und Wasserbewirtschaftung*. 59, Nr. 3, 84–95.

61. GEORGAKAKOS, K. P. (1986): *On the design of national, real-time warning systems with capability for site-specific, flash-flood forecasts*. Bulletin of the American Meteorological Society, 67, Nr. 10, 1233–1239.

62. GEORGAKAKOS, K. P. (2006): *Analytical results for operational flash flood guidance*. Journal of Hydrology, 317, Nr. 1-2, 81–103.

63. GEORGAKAKOS, K. P., A. K. GUETTER and J. A. SPERFSLAGE (1997): *Estimation of flash flood potential for large areas*. IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 239, 87–94.

64. GHONEIM, E. M., N. W. ARNELL and G. M. FOODY (2002): *Characterizing the flash flood hazards potential along the Red Sea coast of Egypt*. In A. SNORASSON, H. P. FINNSDÓTTIR und M. MOSS (Ed.): *The Extremes of the Extremes*. Vol. 271, Wallingford: IAHS, 211–216.

65. GHUMMAN, A. R., Y. M. GHAZAW, A. R. SOHAIL and K. WATANABE (2011): *Runoff forecasting by artificial neural network and conventional model*. Alexandria Engineering Journal, 50, Nr. 4, 345–350.

66. GRIMALDI, S., A. PETROSELLI, E. ARCANGELETTI and F. NARDI (2013): *Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling*. Journal of Hydrology, 487, 39–47.

67. GRIMALDI, S., A. PETROSELLI, F. TAURO and M. PORFIRI (2012): *Time of concentration: A paradox in modern hydrology*. Hydrological Sciences Journal, 57, Nr. 2, 217–228.

68. GRUBBS, F. E. (1969): *Procedures for detecting outlying observations in samples*. Technometrics, 11, Nr. 1, 1–22.

69. GRUNTFEST, E. and J. HANDMER (2001): *Dealing with flash floods: comtemporary issues and future possibilities*. In E. GRUNTFEST und J. HANDMER (Ed.): *Coping with Flash Floods*. Dordrecht: Springer Netherlands, NATO Science Partnership Sub-Series: 2, 3–10.

70. GÜNTNER, A., J. OLSSON, A. CALVER and B. GANNON (2001): *Cascade-based disaggregation of continuous rainfall time series: The influence of climate*. Hydrology and Earth System Sciences, 5, Nr. 2, 145–164.

71. HAPUARACHCHI, H. A. P., Q. J. WANG and T. C. PAGANO (2011): *A review of advances in flash flood forecasting*. Hydrological Processes, 25, Nr. 18, 2771–2784.

72. HENDERSON, F. M. (1966): *Open Channel Flow*. New York, London: Macmillan Publishing Co., Inc., Macmillan Series in civil engineering.

73. HÖLSCHER, J., U. PETRY, M. BERTRAM, M. ANHALT, S. SCHMIDTKE, U. HABERLANDT, H. MÜLLER, S. V.D.HEIJDEN, C. BERNDT, A. VERWORN, M. WALLNER, A. BELLI, J. DIETRICH, G. MEON, K. FÖRSTER, M. GELLESZUN, G. RIEDEL, A. LANGE and F. EGGELSMANN (2012): *Globaler Klimawandel: Wasserwirtschaftliche Folgenabschätzung für das Binnenland.* In *Oberirdische Gewässer Band 33* NLWKN.

74. HORTON, R. E. (1945): *Erosional development of stream and their drainage basins; hydrophysical approach to quantitative morphology*. Geological Society of America Bulletin, 56, Nr. 3, 275–370.

75. Hosking, J. R. and J. R. Wallis (1997): *Regional Frequency Analysis: An Approach based on L-moments*. New York: Cambridge University Press.

76. Hosking, J. R. M. and J. R. Wallis (1993): *Some statistics useful in regional frequency analysis*. Water Resources Research, 29, Nr. 2, 271–281.

77. Huff, F. A. (1967): *Time distribution of rainfall in heavy storms*. Water Resources Research, 3, Nr. 4, 1007–1019.

78. Hunter, N. M., P. D. Bates, M. S. Horritt, A. P. J. de Roo and M. G. F. Werner (2005): *Utility of different data types for calibrating flood inundation models within a GLUE framework*. Hydrology and Earth System Sciences, 9, Nr. 4, 412–430.

79. Jain, V. and R. Sinha (2003): *Derivation of unit hydrograph from GIUH analysis for a Himalayan River*. Water Resources Management, 17, Nr. 5, 355–376.

80. Johnson, R. A. and D. W. Wichern (2002): *Applied Multivariate Statistical Analysis*. 5. Edition. New Jersey, USA: Prentice Hall.

81. Jolliffe, I. T. (2002): *Principal Component Analysis*. 2. Edition. New York: Springer.

82. Jonkman, S. N. and E. Penning-Rowsell (2008): *Human instability in flood flows*. Journal of the American Water Resources Association, 44, Nr. 5, 1208–1218.

83. Jonkman, S. N. and J. K. Vrijling (2008): *Loss of life due to floods*. Journal of Flood Risk Management, 1, Nr. 1, 43–56.

84. Kalantari, Z., A. Nickman, S. W. Lyon, B. Olofsson and L. Folkeson (2014): *A method for mapping flood hazard along roads*. Journal of Environmental Management, 133, 69–77.

85. Karamouz, M., S. Nazif and M. Falahi (2013): *Hydrology and Hydroclimatology: Principles and Applications*. Boca Raton: CRC Press.

86. Karvonen, R. A., A. Hepojoki, J. Kotola, H. K. Huhta and A. Louhio (2000): *The Use of Physical Models in Dam-Break Analysis: RESCDAM Final Report of Helsinki University of Technology (Appendix 2)*. Finland.

87. Kazakis, N., I. Kougias and T. Patsialis (2015): *Assessment of flood hazard areas at a regional scale using an index-based approach and Analytical Hierarchy Process: Application in Rhodope-Evros region, Greece*. The Science of the total environment, 538, 555–563.

88. Keifer, C. J. and H. H. Chu (1957): *Synthetic storm pattern for drainage design*. Journal of the Hydraulics Division, 83, Nr. 4, 1–25.

89. Kelsch, M., E. Caporali and L. G. Lanza (2001): *Hydrometeorology of flash floods*. In E. Gruntfest und J. Handmer (Ed.): *Coping with Flash Floods*. Dordrecht: Springer Netherlands, NATO Science Partnership Sub-Series: 2, 19–35.

90. Khosravi, K., H. R. Pourghasemi, K. Chapi and M. Bahri (2016): *Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: A comparison between Shannon's entropy, statistical index, and weighting factor models*. Environmental monitoring and assessment, 188, Nr. 12, 656.

91. Kim, E. S. and H. I. Choi (2011): *Assessment of vulnerability to extreme flash floods in design storms*. International journal of environmental research and public health, 8, Nr. 7, 2907–2922.

92. Koch, A., G. Meon and C. Leibundgut (2004): *An expert system for the simulation and evaluation of hydrodynamic interaction in floodplains*. In S.-Y. Liong, K.-K. Phoon und V.

BABOVIC (Ed.): *Proceedings of the 6th International Conference on Hydroinformatics*. Singapore: World Scientific, 462–469.

93. KOURGIALAS, N. N. and G. P. KARATZAS (2011): *Flood management and a GIS modelling method to assess flood-hazard areas—a case study*. Hydrological Sciences Journal, 56, Nr. 2, 212–225.

94. KOURGIALAS, N. N. and G. P. KARATZAS (2014): *A hydro-sedimentary modeling system for flash flood propagation and hazard estimation under different agricultural practices*. Natural Hazards and Earth System Science, 14, Nr. 3, 625–634.

95. KOUTROULIS, A. G. and I. K. TSANIS (2010): *A method for estimating flash flood peak discharge in a poorly gauged basin: Case study for the 13–14 January 1994 flood, Giofiros basin, Crete, Greece*. Journal of Hydrology, 385, Nr. 1-4, 150–164.

96. KOUTSOYIANNIS, D. and E. FOUFOULA-GEORGIOU (1993): *A scaling model of a storm hyetograph*. Water Resources Research, 29, Nr. 7, 2345–2361.

97. KRAJEWSKI, W. F. and J. A. SMITH (2002): *Radar hydrology: rainfall estimation*. Advances in Water Resources, 25, Nr. 8-12, 1387–1394.

98. KREYE, P. (2015): *Mesoskalige Bodenwasserhaushaltsmodellierung mit Nutzung von Grundwassermessungen und satellitenbasierten Bodenfeuchtedaten*. PhD dissertation, Technische Universität Braunschweig, Braunschweig.

99. KRON, W. (2002): *Flood risk = hazard x exposure x vulnerability*. In B. WU et al. (Ed.): *Flood defence 2002*. Beijing: Science Press, 82–97.

100. KVOČKA, D., R. A. FALCONER and M. BRAY (2016): *Flood hazard assessment for extreme flood events*. Natural Hazards, 84, Nr. 3, 1569–1599.

101. LE COZANNET, G., M. GARCIN, T. BULTEAU, C. MIRGON, M. L. YATES, M. MÉNDEZ, A. BAILLS, D. IDIER and C. OLIVEROS (2013): *An AHP-derived method for mapping the physical vulnerability of coastal areas at regional scales*. Natural Hazards and Earth System Science, 13, Nr. 5, 1209–1227.

102. LEISMANN, M. and G. MEON (2002): *Das Modell MeadFlow für die praxisgerechte 2D-Modellierung von Strömungen in Flusslandschaften*. Wasserwirtschaft, 92, Nr. 6, 40–44.

103. LEOPOLD, L. B. and T. MADDOCK (1953): *The hydraulic geometry of stream channels and some physiographic implications: Geological survey professional paper 252*. United States Geological Survey.

104. LI, M.-H. and P. CHIBBER (2008): *Overland flow time of concentration on very flat terrains*. Transportation Research Record: Journal of the Transportation Research Board, 2060, 133–140.

105. LIEBSCHER, H. J. (1993): *Hydrology for the water management of large river basins*. Hydrological Sciences Journal, 38, Nr. 1, 1–13.

106. LINSLEY, R. K. (1949): *Applied Hydrology*. New York: McGraw-Hill, Civil engineering series.

107. LOCANTORE, N., J. S. MARRON, D. G. SIMPSON, N. TRIPOLI, J. T. ZHANG, K. L. COHEN, G. BOENTE, R. FRAIMAN, B. BRUMBACK, C. CROUX, J. FAN, A. KNEIP, J. I. MARDEN, D. PEÑA, J. PRIETO, J. O. RAMSAY, M. J. VALDERRAMA and A. M. AGUILERA (1999): *Robust principal component analysis for functional data*. Test, 8, Nr. 1, 1–73.

108. LOHPAISANKRIT, W., G. MEON and T. TINGSANCHALI (2016): *A framework of integrated hydrological and hydrodynamic models using synthetic rainfall for flash flood hazard mapping of*

ungauged catchments in tropical zones. In A. Schumann et al. (Ed.): *The Spatial Dimensions of Water Management - Redistribution of Benefits and Risks*. Vol. 373, Bochum: IAHS, 183–187.

109. Lorenz, M., S. Zeunert, H. Q. Nguyen and G. Meon (2017): *Ökohydrologische Modellierung eines Einzugsgebietes in den sommerfeuchten Tropen im Kontext von Klimawandel und anthropogener Entwicklung*. Hydrologie und Wasserbewirtschaftung, 61, Nr. 6, 408–423.
110. Lumbroso, D. and E. Gaume (2012): *Reducing the uncertainty in indirect estimates of extreme flash flood discharges*. Journal of Hydrology, 414-415, 16–30.
111. LWI-HYWAG and IFW (2015): *Panta Rhei Benutzerhandbuch: Programmdokumentation zur hydrologischen Modellsoftware (unpublished manuscript)*. Braunschweig.
112. Maidment, D. R. (2002): *Arc Hydro: GIS for Water Resources*. 1. Edition. Redlands, Calif.: ESRI Press.
113. Mani, P., C. Chatterjee and R. Kumar (2014): *Flood hazard assessment with multiparameter approach derived from coupled 1D and 2D hydrodynamic flow model*. Natural Hazards, 70, Nr. 2, 1553–1574.
114. Maniak, U. (2010): *Hydrologie und Wasserwirtschaft: Eine Einführung für Ingenieure*. 6. Edition. Berlin and Heidleberg: Springer-Verlag.
115. Marchi, L., M. Borga, E. Preciso and E. Gaume (2010): *Characterisation of selected extreme flash floods in Europe and implications for flood risk management*. Journal of Hydrology, 394, Nr. 1-2, 118–133.
116. Marchi, L., M. Borga, E. Preciso, M. Sangati, E. Gaume, V. Bain, G. Delrieu, L. Bonnifait and N. Pogacnik (2009): *Comprehensive post-event survey of a flash flood in Western Slovenia: Observation strategy and lessons learned*. Hydrological Processes, 23, Nr. 26, 3761–3770.
117. Margane, A. and T. Tattong (1999): *Aspects of the hydrogeology of the Chiang Mai-Lamphun basin, Thailand that are important for groundwater management*. Zeitschrift für Angewandte Geologie, 45, 188–197.
118. Marshall, J. S. and W. M. K. Palmer (1948): *The distribution of raindrops with size*. Journal of Meteorology, 5, Nr. 4, 165–166.
119. Martina, M. L. V., E. Todini and A. Libralon (2006): *A Bayesian decision approach to rainfall thresholds based flood warning*. Hydrology and Earth System Sciences, 10, Nr. 3, 413–426.
120. Masoud, A. A. (2004): *Flash flood potential, mitigation, and floodwater resource management integrating remote sensing and GIS technologies in Safaga Area, Egypt*. Journal of Geosciences, Osaka City University, 47, 21–38.
121. McKay, M. D., R. J. Beckman and W. J. Conover (1979): *A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code*. Technometrics, 21, Nr. 2, 239.
122. Meon, G. and N. Phuoc (2014): *BMBF-VNU-joint research project “Environmental and Water Protection Technologies of Coastal Zones in Vietnam" (EWATEC-COAST)*. In G. Meon et al. (Ed.): *EWATEC-COAST: Technologies for environmental and water protection of coastal regions in Vietnam*. Göttingen: Cuvillier, HYWAG-Schriftenreihe, 1–15.

123. MEON, G., G. RIEDEL, M. ANHALT and J. HÖLSCHER (2015): *Innovative Hochwasservorhersage für kleine und mittlere Einzugsgebiete*. In M. EVERS und B. DIEKKRÜGER (Ed.): *Aktuelle Herausforderungen im Flussgebiets- und Hochwassermanagement*. Bonn, Forum für Hydrologie und Wasserbewirtschaftung.
124. MERZ, R. and G. BLÖSCHL (2005): *Flood frequency regionalisation – spatial proximity vs. catchment attributes*. Journal of Hydrology, 302, Nr. 1-4, 283–306.
125. MINEA, G. (2013): *Assessment of the flash flood potential of Bâsca River Catchment (Romania) based on physiographic factors*. Central European Journal of Geosciences, 5, Nr. 3, 344–353.
126. MODRICK, T. M. and K. P. GEORGAKAKOS (2014): *Regional bankfull geometry relationships for southern California mountain streams and hydrologic applications*. Geomorphology, 221, 242–260.
127. MOLNAR, P. and P. BURLANDO (2005): *Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model*. Atmospheric Research, 77, Nr. 1-4, 137–151.
128. MONTEITH, J. L. (1973): *Principles of Environmental Physics*. London: Edward Arnold, Series of student texts in contemporary biology.
129. MONTERO, R. A., D. SCHWANENBERG, M. HATZ and M. BRINKMANN (2013): *Simplified hydraulic modelling in model predictive control of flood mitigation measures along rivers*. Journal of Applied Water Engineering and Research, 1, Nr. 1, 17–27.
130. MOSLEY, M. P. and A. I. MCKERCHAR (1993): *Streamflow*. In D. R. MAIDMENT (Ed.): *Handbook of Hydrology*. New York: McGraw-Hill, 8.1–8.39.
131. MULVANEY, T. J. (1851): *On the use of self-registering rain and flood gauges in making observations of the relations of rain fall and flood discharges in a given catchment*. Transactions of the Institution of Civil Engineers of Ireland, 4, 18–33.
132. NANDI, A., A. MANDAL, M. WILSON and D. SMITH (2016): *Flood hazard mapping in Jamaica using principal component analysis and logistic regression*. Environmental Earth Sciences 75, Nr. 6.
133. NASH, J. E. and J. V. SUTCLIFFE (1970): *River flow forecasting through conceptual models part I – A discussion of principles*. Journal of Hydrology, 10, Nr. 3, 282–290.
134. NOAA (2010): *Flash Flood Early Warning System Reference Guide*. University Corporation for Atmospheric Research.
135. NORBIATO, D., M. BORGA, S. DEGLI ESPOSTI, E. GAUME and S. ANQUETIN (2008): *Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins*. Journal of Hydrology, 362, Nr. 3-4, 274–290.
136. NORBIATO, D., M. BORGA, M. SANGATI and F. ZANON (2007): *Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood*. Journal of Hydrology, 345, Nr. 3-4, 149–166.
137. NRC (2005): *Flash Flood Forecasting over Complex Terrain: With an Assessment of the Sulphur Mountain NEXRAD in Southern California*. Washington, D.C.: National Academies Press, National Research Council (U.S.).

138. OGDEN, F. L., H. O. SHARIF, S. SENARATH, J. A. SMITH, M. L. BAECK and J. R. RICHARDSON (2000): *Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997*. Journal of Hydrology, 228, Nr. 1-2, 82–100.

139. OMRAN, A., D. SCHRÖDER, A. EL-RAYES and M. GERIESH (2011): *Flood hazard assessment in Wadi Dahab, Egypt based on basin morphometry using GIS techniques*. In A. CAR, G. GRIESEBNER und J. STROBL (Ed.): *Geospatial crossroads GI_Forum '11*. Berlin: Herbert Wichmann Verlag, 1–11.

140. ONOF, C., R. E. CHANDLER, A. KAKOU, P. NORTHROP, H. S. WHEATER and V. ISHAM (2000): *Rainfall modelling using Poisson-cluster processes: a review of developments*. Stochastic Environmental Research and Risk Assessment, 14, Nr. 6, 384–411.

141. PALLARD, B., A. CASTELLARIN and A. MONTANARI (2009): *A look at the links between drainage density and flood statistics*. Hydrology and Earth System Sciences, 13, Nr. 7, 1019–1029.

142. PASCHE, E., S. KRÄSSIG, K. LIPPERT, H. NASERMOADDELI, W. PLÖGER and S. RATH (2006): *Wie viel Physik braucht die Strömungsberechnung in der Ingenieurpraxis?* In H. HORLACHER und K. GRAW (Ed.): *Strömungssimulation im Wasserbau*. Dresden, Dresdner wasserbauliche Mitteilungen, 237–260.

143. PATTON, P. C. and V. R. BAKER (1976): *Morphometry and floods in small drainage basins subject to diverse hydrogeomorphic controls*. Water Resources Research, 12, Nr. 5, 941–952.

144. PEKÁROVÁ, P., A. SVOBODA, P. MIKLÁNEK, P. ŠKODA, D. HALMOVÁ and J. PEKÁR (2012): *Estimating flash flood peak discharge in Gidra and Parná basin: Case study for the 7-8 June 2011 flood*. Journal of Hydrology and Hydromechanics, 60, Nr. 3, 206–216.

145. PENMAN, H. L. (1948): *Natural evaporation from open water, bare soil and grass*. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 193, Nr. 1032, 120–145.

146. PENNING-ROWSELL, E., P. FLOYD, D. RAMSBOTTOM and S. SURENDRAN (2005): *Estimating injury and loss of life in floods: A deterministic framework*. Natural Hazards, 36, Nr. 1-2, 43–64.

147. PILGRIM, D. H. and I. CORDERY (1993): *Food runoff*. In D. R. MAIDMENT (Ed.): *Handbook of Hydrology*. New York: McGraw-Hill, 9.1–9.42.

148. PIOTROWSKI, A., J. J. NAPIÓRKOWSKI and P. M. ROWIŃSKI (2006): *Flash-flood forecasting by means of neural networks and nearest neighbour approach – a comparative study*. Nonlinear Processes in Geophysics, 13, Nr. 4, 443–448.

149. PONCE, V. M. (1989): *Engineering Hydrology*. Englewood Cliffs: Prentice Hall.

150. PRIESTLEY, C. H. B. and R. J. TAYLOR (1972): *On the assessment of surface heat flux and evaporation using large-scale parameters*. Monthly Weather Review, 100, Nr. 2, 81–92.

151. PUI, A., A. SHARMA, R. MEHROTRA, B. SIVAKUMAR and E. JEREMIAH (2012): *A comparison of alternatives for daily to sub-daily rainfall disaggregation*. Journal of Hydrology, 470-471, 138–157.

152. R CORE TEAM (2015): *R: A Language and Environment for Statistical Computing*. ⟨URL: `https://www.R-project.org/`⟩.

153. RADICE, A., E. GIORGETTI, D. BRAMBILLA, L. LONGONI and M. PAPINI (2012): *On integrated sediment transport modelling for flash events in mountain environments*. Acta Geophysica, 60, Nr. 1, 191–213.

154. RALLISON, M. C. and M. WERNER (1981): *Past, present, and future SCS runoff procedure*. In V. P. SINGH (Ed.): *Rainfall-Runoff Relationship*. Littleton: Water Resources Publications, 353–364.

155. RAMSBOTTOM, D., P. FLOYD and E. PENNING-ROWSELL (2003): *Flood Risks to People: Phase 1, R&D Technical Report FD2317/TR*. London.

156. RAMSBOTTOM, D., S. WADE, V. BAIN, P. FLOYD, E. PENNING-ROWSELL, T. WILSON, A. FERNANDEZ, M. HOUSE and S. SURENDRAN (2006): *Flood Risks to People: Phase 2, FD2321/TR1 the Flood Risks to People Methodology*. London.

157. REED, S., D. JOHNSON and T. SWEENEY (2002): *Application and national geographic information system database to support two-year flood and threshold runoff estimates*. Journal of Hydrologic Engineering, 7, Nr. 3, 209–219.

158. REED, S., J. SCHAAKE and Z. ZHANG (2007): *A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations*. Journal of Hydrology, 337, Nr. 3-4, 402–420.

159. RENCHER, A. C. (2002): *Methods of Multivariate Analysis*. 2. Edition. New York: J. Wiley, Wiley series in probability and mathematical statistics.

160. RIEDEL, G., M. ANHALT, S. MEYER, E. WEIGL and G. MEON (2017): *Erfahrung mit Radarprodukten bei der operationellen Hochwasservorhersage in Niedersachsen*. KW - Korrespondenz Wasserwirtschaft, 11, 664–671.

161. RIEDEL, G. (2004): *Ein hydrologisches Modell für tidebeeinflusste Flussgebiete*. PhD dissertation, Technische Universität Braunschweig, Braunschweig.

162. RITTER, D. F. (1978): *Process Geomorphology*. Dubuque, Iowa: W.C. Brown Co.

163. RODRÍGUEZ-ITURBE, I., D. R. COX and V. ISHAM (1987): *Some models for rainfall based on stochastic point processes*. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 410, Nr. 1839, 269–288.

164. RODRÍGUEZ-ITURBE, I. and J. B. VALDÉS (1979): *The geomorphologic structure of hydrologic response*. Water Resources Research, 15, Nr. 6, 1409–1420.

165. ROSBJERG, D., G. BLÖSCHL, D. BURN, A. CASTELLARIN, B. CROKE, G. DI BALDASSARRE, V. IACOBELLIS, T. KJELDSEN, G. KUCZERA, R. MERZ, A. MONTANARI, D. MORRIS, T. OUARDA, L. REN, M. ROGGER, J. SALINAS, E. TOTH and A. VIGLIONE (2013): *Prediction of floods in ungauged basins*. In G. BLÖSCHL et al. (Ed.): *Runoff Prediction in Ungauged Basins*. Cambridge University Press, 189–226.

166. ROUX, H., D. LABAT, P. A. GARAMBOIS, M.-M. MAUBOURGUET, J. CHORDA and D. DARTUS (2011): *A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments*. Natural Hazards and Earth System Science, 11, Nr. 9, 2567–2582.

167. ROZALIS, S., E. MORIN, Y. YAIR and C. PRICE (2010): *Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions*. Journal of Hydrology, 394, Nr. 1-2, 245–255.

168. RUIZ-VILLANUEVA, V., J. M. BODOQUE, A. DÍEZ-HERRERO, M. A. EGUIBAR and E. PARDO-IGÚZQUIZA (2013): *Reconstruction of a flash flood with large wood transport and its influence on hazard patterns in an ungauged mountain basin*. Hydrological Processes, 27, Nr. 24, 3424–3437.
169. RULLI, M. C. and R. ROSSO (2002): *An integrated simulation method for flash-flood risk assessment: 1. Frequency predictions in the Bisagno River by combining stochastic and deterministic methods*. Hydrology and Earth System Sciences, 6, Nr. 2, 267–284.
170. RUTTER, A. J., K. A. KERSHAW, P. C. ROBINS and A. J. MORTON (1971): *A predictive model of rainfall interception in forests, 1. derivation of the model from observations in a plantation of corsican pine*. Agricultural Meteorology, 9, 367–384.
171. SCHMITZ, G. H. and J. CULLMANN (2008): *PAI-OFF: A new proposal for online flood forecasting in flash flood prone catchments*. Journal of Hydrology, 360, Nr. 1-4, 1–14.
172. SCHUMACHER, R. S. and R. H. JOHNSON (2006): *Characteristics of U.S. extreme rain events during 1999–2003*. Weather and Forecasting, 21, Nr. 1, 69–85.
173. SCHUMM, S. A. (1956): *Evolution of drainage systems and slopes in Badlands at Perth Amboy, New Jersey*. Geological Society of America Bulletin, 67, Nr. 5, 597–646.
174. SCHWARZ, G. (1978): *Estimating the dimension of a model*. The Annals of Statistics, 6, Nr. 2, 461–464.
175. SCOTT, K. M. and G. C. GRAVLEE (1968): *Flood Surge on the Rubicon River, California, Hydrology, Hydraulics, and Boulder Transport: Geological Survey Professional Paper 422-M*. United States Government Printing Office.
176. SENE, K. (2013): *Flash floods: Forecasting and Warning*. Dordrecht and New York: Springer.
177. SHARMA, A. and R. MEHROTRA (2010): *Rainfall generation*. In F. Y. TESTIK und M. GEBREMICHAEL (Ed.): *Rainfall*. Washington, DC: American Geophysical Union, Geophysical monograph, 215–246.
178. SHAW, E. M., K. J. BEVEN, N. A. CHAPPELL and R. LAMB (2011): *Hydrology in Practice*. 4. Edition. New York: Spon Press.
179. SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. London and New York: Chapman and Hall, Monographs on statistics and applied probability.
180. SMITH, G. (2003): *Flash flood potential: Determining the hydrologic response of FFMP basins to heavy rain by analyzing their physiographic characteristics*. ⟨URL: `http://www.cbrfc.noaa.gov/papers/ffp_wpap.pdf`⟩.
181. SMITH, K. and R. WARD (1998): *Floods: Physical Processes and Human Impacts*. New York: Wiley.
182. SNYDER, F. F. (1938): *Synthetic unit-graphs*. Transactions, American Geophysical Union, 19, Nr. 1, 447.
183. SØRENSEN, R., U. ZINKO and J. SEIBERT (2006): *On the calculation of the topographic wetness index: Evaluation of different methods based on field observations*. Hydrology and Earth System Sciences, 10, Nr. 1, 101–112.
184. STEFFLER, P. and J. BLACKBURN (2002): *River2D: Two-dimensional depth averaged model of river hydrodynamics and fish habitat: Introduction to depth averaged modeling and user's manual*. ⟨URL: `http://www.river2d.ualberta.ca/`⟩.

185. Stödter, A. (1994): *GIS-gestützte Ermittlung von Abflußkonzentrationsparametern für ein konzeptionelles Hochwassermodell*. PhD dissertation, Technische Universität Braunschweig, Braunschweig.
186. Strahler, A. N. (1952): *Dynamic basis of geomorphology*. Geological Society of America Bulletin, 63, Nr. 9, 923.
187. Strahler, A. N. (1957): *Quantitative analysis of watershed geomorphology*. Transactions, American Geophysical Union, 38, Nr. 6, 913.
188. Sweeney, T. (1992): *Modernized Areal Flash Flood Guidance*. Vol. 44, NOAA technical memorandum NWS HYDRO. Springfield, VA.
189. Tarboton, D. (2003): *Rainfall-runoff processes*. ⟨URL: `http://hydrology.usu.edu/`⟩.
190. Tehrany, M. S., B. Pradhan and M. N. Jebur (2013): *Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS*. Journal of Hydrology, 504, 69–79.
191. Tehrany, M. S., B. Pradhan and M. N. Jebur (2014): *Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS*. Journal of Hydrology, 512, 332–343.
192. Thompson, S. A. (1999): *Hydrology for Water Management*. Rotterdam, Netherlands and Brookfield, VT, USA: A.A. Balkema.
193. Tingsanchali, T. and M. F. Karim (2005): *Flood hazard and risk analysis in the southwest region of Bangladesh*. Hydrological Processes, 19, Nr. 10, 2055–2069.
194. Tou, J. T. and R. C. Gonzalez (1981): *Pattern Recognition Principles*. Vol. 7, Applied mathematics and computation. 4. Edition. Reading, Massachusetts: Addison-Wesley.
195. Toukourou, M., A. Johannet, G. Dreyfus and P.-A. Ayral (2011): *Rainfall-runoff modeling of flash floods in the absence of rainfall forecasts: the case of "Cévenol flash floods"*. Applied Intelligence, 35, Nr. 2, 178–189.
196. USDA (1986): *Urban Hydrology for Small Watersheds: TR-55*. United States Department of Agriculture.
197. Van Ngo, T. T. (2000): *Abflußkonzentrationszeit von Sturzhochwasser (Flash Floods) in Monsunregionen*. PhD dissertation, Technische Universität Braunschweig, Braunschweig.
198. Van Rijsbergen, C. J. (1975): *Information Retrieval*. London: Butterworths.
199. Velasco-Forero, C., C. Corral and D. Sempere-Torres (2006): *Parameterisation of hydrological model for application in ungauged basins*. FLOODsite Consortium.
200. Veneziano, D. and P. Villani (1999): *Best linear unbiased design hyetograph*. Water Resources Research, 35, Nr. 9, 2725–2738.
201. Viglione, A. (2014): *nsRFA: Non-supervised Regional Frequency Analysis*. ⟨URL: `https://CRAN.R-project.org/package=nsRFA`⟩.
202. Wagener, T., G. Blöschl, D. C. Goodrich, H. V. Gupta, M. Sivapalan, Y. Tachikawa, P. A. Troch and M. Weiler (2013): *A synthesis framework for runoff prediction in ungauged basins*. In G. Blöschl et al. (Ed.): *Runoff Prediction in Ungauged Basins*. Cambridge University Press, 11–28.

203. Wahid, A., M. Madden, F. Khalaf and I. Fathy (2016): *Geospatial analysis for the determination of hydro-morphological characteristics and assessment of flash flood potentiality in arid coastal plains: A case in southwestern Sinai, Egypt*. Earth Sciences Research Journal, 20, Nr. 1, E1–E9.

204. Wang, C. T., V. K. Gupta and E. Waymire (1981): *A geomorphologic synthesis of nonlinearity in surface runoff*. Water Resources Research, 17, Nr. 3, 545–554.

205. Wang, Z., C. Lai, X. Chen, B. Yang, S. Zhao and X. Bai (2015): *Flood hazard risk assessment model based on random forest*. Journal of Hydrology, 527, 1130–1141.

206. Webb, R. H. and R. D. Jarrett (2002): *One-Dimensional estimation techniques for discharges of paleofloods and historical floods*. In P. K. House et al. (Ed.): *Ancient Floods, Modern Hazards: Principles and Applications of Paleoflood Hydrology*. Washington, D. C.: American Geophysical Union, 111–125.

207. Wheater, H. S., R. E. Chandler, C. J. Onof, V. S. Isham, E. Bellone, C. Yang, D. Lekkas, G. Lourmas and M.-L. Segond (2005): *Spatial-temporal rainfall modelling for flood risk estimation*. Stochastic Environmental Research and Risk Assessment, 19, Nr. 6, 403–416.

208. WMO (1994): *Guide to Hydrological Practices*. World Meteorological Organization.

209. Wood, E. F. and C. S. Hebson (1986): *On hydrologic similarity: 1. derivation of the dimensionless flood frequency curve*. Water Resources Research, 22, Nr. 11, 1549–1554.

210. Wood, S. H. and A. D. Ziegler (2008): *Floodplain sediment from a 100-year-recurrence flood in 2005 of the Ping River in northern Thailand*. Hydrology and Earth System Sciences, 12, Nr. 4, 959–973.

211. Woods, R. A. and M. Sivapalan (1997): *A connection between topographically driven runoff generation and channel network structure*. Water Resources Research, 33, Nr. 12, 2939–2950.

212. Xia, J., R. A. Falconer, B. Lin and G. Tan (2011): *Numerical assessment of flood hazard risk to people and vehicles in flash floods*. Environmental Modelling & Software, 26, Nr. 8, 987–998.

213. Youssef, A. M., B. Pradhan and S. A. Sefry (2016): *Flash flood susceptibility assessment in Jeddah city (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models*. Environmental Earth Sciences, 75, Nr. 1, 23.

214. Youssef, A. M., B. Pradhan and A. M. Hassan (2011): *Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery*. Environmental Earth Sciences, 62, Nr. 3, 611–623.

215. Yucel, I. (2015): *Assessment of a flash flood event using different precipitation datasets*. Natural Hazards, 79, Nr. 3, 1889–1911.

216. Zaharia, L., G. Minea, G. Ioana-Toroimac, R. Barbu and I. Sârbu (2012): *Estimation of the Areas with Accelerated Surface Runoff in the upper Prahova watershed (Romanian Carpathians)*. In *Balwois 2012 International Conference on Water, Climate and Environment.*, 1–10.

217. Zemzami, M., L. Benaabidate, B. Layan and A. Dridri (2013): *Design flood estimation in ungauged catchments and statistical characterization using principal components analysis: Application of Gradex method in Upper Moulouya*. Hydrological Processes, 27, Nr. 2, 186–195.

218. Zeng, Z., G.-G. Tang, D. Long, H. Xu, Y. Chen and Y. Hong (2015): *Development of GIS-based FFPI for China's flash flood forecasting*. In G. Xu und D. Richardson (Ed.): *Proceedings of the 23rd International Conference on Geoinformatics*. Wuhan, China, 1–4.
219. Zhang, D.-W., J. Quan, H.-B. Zhang, F. Wang, H. Wang and X.-Y. He (2015): *Flash flood hazard mapping: A pilot case study in Xiapu River Basin, China*. Water Science and Engineering, 8, Nr. 3, 195–204.
220. Zhang, W. and D. R. Montgomery (1994): *Digital elevation model grid size, landscape representation, and hydrologic simulations*. Water Resources Research, 30, Nr. 4, 1019–1028.
221. Zienkiewicz, O. C., R. L. Taylor and P. Nithiarasu (2014): *The Finite Element Method for Fluid Dynamics*. 7. Edition. Oxford and Waltham, Mass.: Butterworth-Heinemann.

[This page intentionally left blank]

A Appendix

A.1 Data

A.1.1 List of meteorological stations

Table A.1: List of meteorological stations with daily data of meteorological variables: P=rainfall; T=temperature; S=sunshine duration; p=relative humidity; e=evaporation; and W=wind speed

No.	Code	Station name	Latitude (°) N	Longitude (°) E	Operator	Measurements
1	090201*	Nam Pua at Ban Na Fang	19.22	100.97	DWR	P
2	331003*	Pua	19.18	100.92	TMD	P
3	331005*	Chiang Klang	19.29	100.87	TMD	P
4	331006*	Mae Charim	18.73	101.02	TMD	P
5	331009*	Song Khwae	19.36	100.71	TMD	P
6	331012	Chaloem Phrakiat	19.58	101.08	TMD	P
7	331014	Santisuk	19.08	100.92	TMD	P
8	331201*	Muang Nan	18.78	100.78	TMD	P**, T, S, p, e, W
9	331301*	Nan Agrometeorological Station	18.87	100.75	TMD	P, T, S, p, e, W
10	331401*	Tha Wang Pha	19.11	100.80	TMD	P, T, p, e
11	331402	Thung Chang	19.41	100.89	TMD	P, T, p, e

* Station contains more than 18 years of rainfall data.
** 3-hourly rainfall records.

Table A.2: List of 15-min rain gauges used in the upper part of Nan River basin and their available periods

No.	Code	Station name	Latitude (°) N	Longitude (°) E	Available period
1	DWR0020	Muang Nan	100.63	18.92	2006-2012
2	DWR0033	Bo Kluea	101.18	19.31	2006-2012
3	DWR0040	Bo Kluea	101.18	19.28	2006-2012
4	DWR0056	Chaloem Phrakiat	101.02	19.57	2007-2012
5	DWR0069	Tha Wang Pha	100.93	19.06	2006-2012
6	DWR0072	Thung Chang	100.83	19.45	2006-2012
7	DWR0076	Bo Kluea	101.19	19.23	2008-2012
8	DWR0088	Song Khwae	100.68	19.43	2006-2012
9	DWR0089	Song Khwae	100.65	19.37	2007-2012
10	DWR0169	Pua	101.01	19.15	2008-2012
11	DWR0181	Pua	101.00	19.12	2009-2012
12	DWR0185	Chiang Klang	100.81	19.33	2009-2012
13	DWR0239	Tha Wang Pha	100.75	19.02	2009-2012
14	DWR0336	Chaloem Phrakiat	101.06	19.58	2009-2012
15	DWR0337	Chiang Klang	100.95	19.28	2009-2012
16	DWR0338	Thung Chang	100.93	19.52	2009-2012
17	DWR0339	Thung Chang	100.90	19.49	2009-2012
18	DWR0340	Thung Chang	100.83	19.39	2009-2012
19	DWR0342	Song Khwae	100.60	19.38	2009-2012
20	DWR0348	Bo Kluea	101.15	19.15	2009-2012
21	DWR0351	Thung Chang	100.97	19.39	2009-2012
22	DWR0352	Tha Wang Pha	100.79	19.14	2009-2012
23	DWR0353	Tha Wang Pha	100.87	19.07	2009-2012
24	DWR0447	Muang Nan	100.55	18.85	2010-2012
25	DWR0541	Pua	101.07	18.99	2010-2012
26	DWR0542	Mae Charim	101.03	18.79	2010-2014

Note: DWR stands for the Department of Water Resources

A.2 Methods

A.2.1 Efficiency criteria

The efficiency criteria described in this section are used to evaluate the simulation results with observations. In this dissertation, Nash-Sutcliffe efficiency index, relative peak error, coefficient of determination, root mean square error, mean absolute error, and percent bias were used.

- Nash-Sutcliffe Efficiency Index (*EI*): this statistical criterion introduced by Nash and Sutcliffe (1970) is commonly used to quantitatively describe the ability of model to reproduce the observed data (e.g., discharges). It is computed as follows:

  $$EI = 1 - \frac{\sum_{i=1}^{n} \left(X_{obs,i} - X_{sim,i}\right)^2}{\sum_{i=1}^{n} \left(X_{obs,i} - \overline{X}_{sim,i}\right)^2} \tag{A.1}$$

  where

$X_{obs,i}$	is the observed variable at time *i* [e.g., m^3 s^{-1}];
$X_{sim,i}$	is the simulated variable at time *i* [e.g., m^3 s^{-1}];
$\overline{X}_{obs}$	is the average of the observed variable [e.g., m^3 s^{-1}];
$\overline{X}_{sim}$	is the average of the simulated variable [e.g., m^3 s^{-1}];
n	is the number of time step;

  The range of *EI* varies between negative infinity and one, where an *EI* of one corresponds to a perfect match between the simulated data (X_{sim}) and observed data (X_{obs}). An *EI* of zero indicates that the model predictions are as accurate as the mean of the observed data ($\overline{X}_{obs}$), whereas an *EI* of lower than zero indicates that the mean value of observed data is a better predictor than the model.

- Relative Peak Error (*RPE*): this index measures the relative difference between observed and simulated peaks. It is computed as follows:

  $$RPE = \frac{X_{obs} - X_{sim}}{X_{obs}} \times 100\% \tag{A.2}$$

  A positive value of *RPE* indicates that the observed peak is greater than the simulated peak, and vice versa.

- Coefficient of determination (r^2): this index measures the strength of a linear relationship between simulated and observed values. It is computed as follows:

  $$r^2 = \left[\frac{\sum_{i=1}^{n} \left(X_{obs,i} - \overline{X}_{obs}\right)\left(X_{sim,i} - \overline{X}_{sim}\right)}{\sqrt{\sum_{i=1}^{n} \left(X_{obs,i} - \overline{X}_{obs}\right)^2 \sum_{i=1}^{n} \left(X_{sim,i} - \overline{X}_{sim}\right)^2}}\right]^2 \tag{A.3}$$

  The range of r^2 lies between zero and one, which describes how well observed data are replicated by the model. A value of zero indicates no correlation at all, whereas a value of one indicates a perfect match between both sets of data.

- Root Mean Square Error (*RMSE*): this index measures individual differences (called residuals) between the observed and simulated values. It is computed as follows:

  $$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(X_{obs,i} - X_{sim,i}\right)^2}{n}} \tag{A.4}$$

  Smaller values of *RMSE* indicate better agreement between the observed and simulated values.

- Mean Absolute Error (*MAE*): this index measures how far simulated values are away from observed values. It is computed as follows:

  $$MAE = \frac{\sum_{i}^{n} \mid X_{sim,i} - X_{obs,i} \mid}{n} \tag{A.5}$$

  Similarly to *RMSE*, smaller values of *MAE* indicate better agreement between the observed and simulated values.

- Percent bias (*PBIAS*): this index measures the average tendency between observed and simulated values. It is calculated as follows:

  $$PBIAS = \frac{\sum_{i=1}^{n} \left(X_{sim,i} - X_{obs,i}\right)}{\sum_{i=1}^{n} X_{obs}} \times 100\% \tag{A.6}$$

  This index is expressed as percentage. Absence of bias corresponds to zero. Negative values of *PBIAS* indicate model underestimation bias, whereas positive values indicate model overestimation bias.

A.3 Results

A.3.1 Rainfall depth-duration-frequency curves

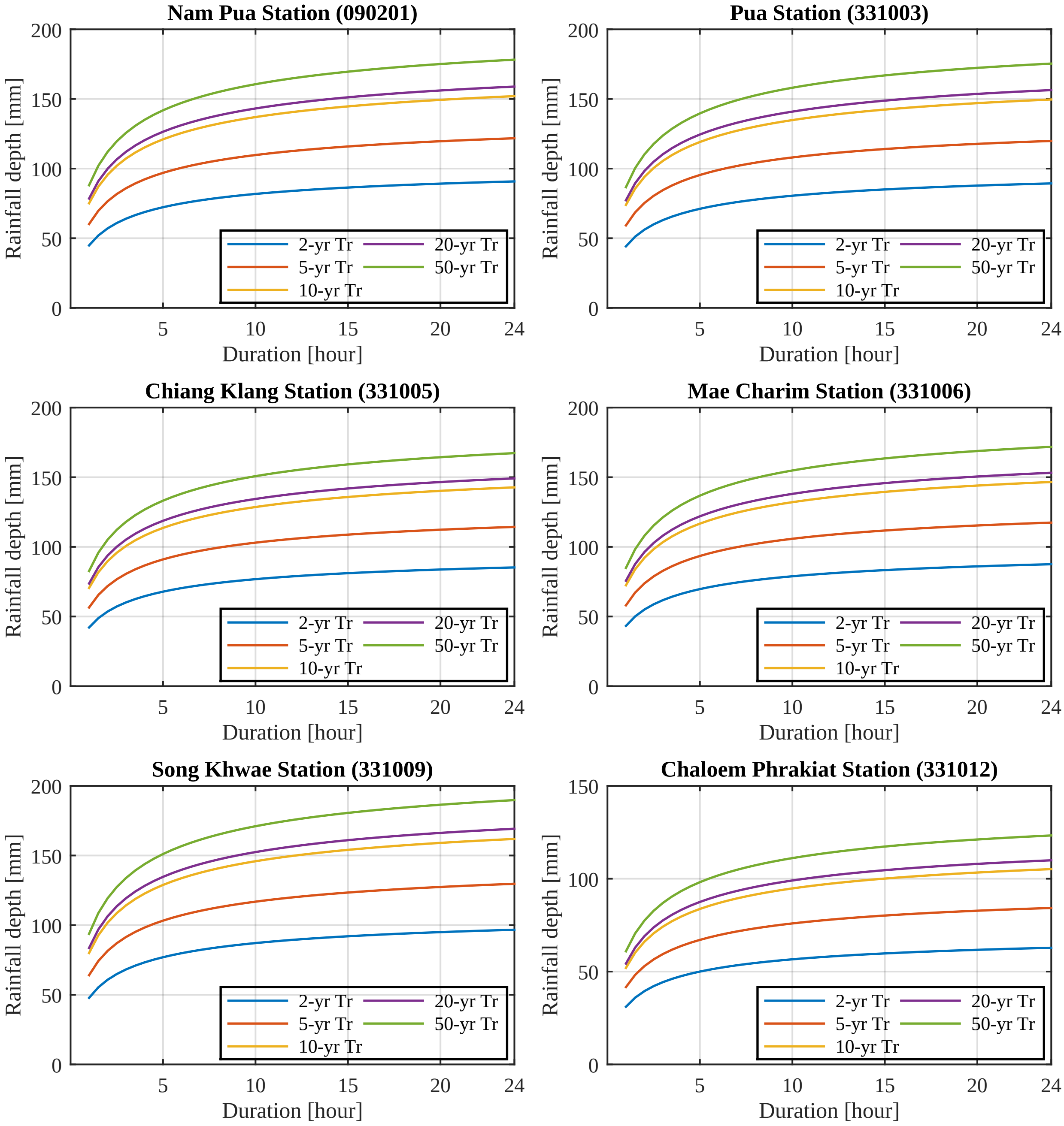


Figure A.1: Rainfall depth-duration-frequency curves.

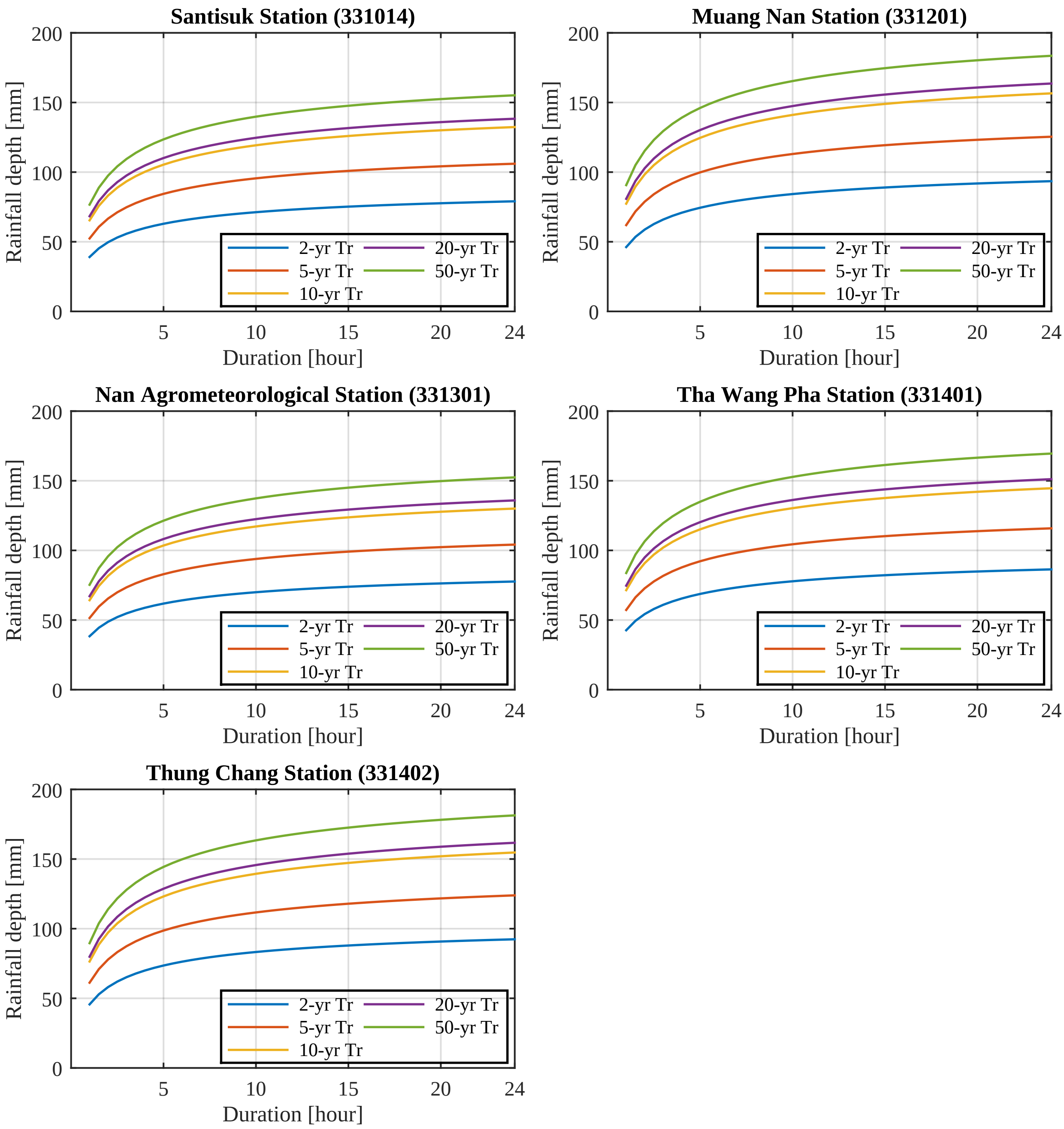


Figure A.1: Rainfall depth-duration-frequency curves (continued).

A.3.2 Principal component analysis

Table A.3: All principal components and their variance for the Expt. 3 (physiographic variables and IDW-rainfall spatial distribution derived from the three rain gauges under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.09	**0.95**	0.28	0.00	-0.04	-0.09	0.10
Rbuffer	0.48	0.07	-0.47	0.26	-0.04	**-0.58**	-0.38
Surface slope	0.15	-0.20	**0.57**	**0.77**	-0.14	-0.02	0.00
Elevation	**0.52**	0.07	-0.01	-0.05	0.05	**0.72**	-0.45
TWI	0.04	0.00	0.08	0.10	**0.99**	-0.07	0.03
FL	**0.57**	-0.08	-0.19	0.00	-0.03	0.10	**0.79**
IDW-rainfall	-0.38	0.23	**-0.57**	**0.58**	0.03	0.36	0.12
Variance [%]	51.36	17.61	15.19	7.82	5.15	2.42	0.45
Cumulative variance [%]	51.36	68.97	84.16	91.98	97.14	99.55	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Table A.4: All principal components and their variance for the Expt. 4 (physiographic variables and KG-rainfall spatial distribution derived from the three rain gauges under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.09	**0.98**	0.07	-0.04	-0.02	-0.11	0.10
Rbuffer	0.49	-0.02	-0.33	-0.05	**0.60**	-0.43	-0.32
Surface slope	0.13	-0.09	**0.91**	-0.15	0.35	-0.04	0.00
Elevation	**0.50**	0.08	0.03	0.07	-0.15	**0.69**	-0.49
TWI	0.04	0.02	0.13	**0.98**	0.06	-0.10	0.03
FL	**0.56**	-0.10	-0.10	-0.02	0.03	0.18	**0.79**
KG-rainfall	-0.41	0.11	-0.17	0.04	**0.70**	**0.53**	0.13
Variance [%]	55.69	18.19	12.32	5.35	4.92	3.04	0.49
Cumulative variance [%]	55.69	73.87	86.19	91.54	96.47	99.51	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Table A.5: All principal components and their variance for the Expt. 5 (physiographic variables and TP-rainfall spatial distribution derived from the three rain gauges under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	-0.03	0.12	**0.98**	0.08	-0.04	0.07	0.11
Rbuffer	0.22	**0.53**	-0.06	-0.10	-0.05	**0.72**	-0.38
Surface slope	-0.19	0.15	-0.11	**0.95**	-0.14	0.06	-0.01
Elevation	-0.02	**0.55**	0.04	-0.04	0.05	**-0.66**	**-0.50**
TWI	-0.03	0.05	0.01	0.13	**0.99**	0.08	0.04
FL	0.05	**0.61**	-0.15	-0.10	-0.03	-0.08	**0.77**
TP-rainfall	**0.95**	-0.11	0.03	0.22	0.01	-0.16	0.04
Variance [%]	36.95	35.60	13.52	7.39	4.03	2.14	0.38
Cumulative variance [%]	36.95	72.54	86.06	93.45	97.48	99.62	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Table A.6: All principal components and their variance for the Expt. 6 (physiographic variables and IDW-rainfall spatial distribution derived from the two rain gauges under the R250611 scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	-0.06	**0.98**	0.09	-0.03	-0.04	-0.13	0.09
Rbuffer	-0.44	0.02	-0.38	**0.66**	-0.03	-0.31	-0.36
Surface slope	-0.12	-0.09	**0.88**	0.42	-0.14	-0.01	-0.01
Elevation	-0.47	0.10	0.03	-0.17	0.05	**0.75**	-0.41
TWI	-0.04	0.02	0.12	0.07	**0.99**	-0.07	0.03
FL	**-0.52**	-0.07	-0.12	0.11	-0.02	0.16	**0.82**
IDW-rainfall	**0.54**	0.13	-0.18	**0.59**	0.03	**0.54**	0.14
Variance [%]	59.24	16.82	11.36	5.78	4.90	1.45	0.46
Cumulative variance [%]	59.24	76.06	87.41	93.19	98.09	99.54	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Table A.7: All principal components and their variance for the Expt. 7 (physiographic variables and IDW-rainfall spatial distribution derived from the three rain gauges under the R$_{assumed}$ scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.07	**0.98**	0.04	-0.04	0.03	-0.12	0.07
Rbuffer	0.47	-0.02	-0.29	-0.06	**0.77**	0.04	-0.31
Surface slope	0.11	-0.06	**0.94**	-0.16	0.26	0.02	-0.01
Elevation	0.48	0.09	0.06	0.07	-0.43	**0.68**	-0.33
TWI	0.03	0.02	0.13	**0.98**	0.12	-0.06	0.02
FL	**0.54**	-0.09	-0.06	-0.02	-0.02	0.03	**0.84**
IDW-rainfall	-0.49	0.09	-0.04	0.02	0.37	**0.72**	0.31
Variance [%]	59.80	17.66	11.68	5.19	4.24	0.96	0.47
Cumulative variance [%]	59.80	77.46	89.14	94.33	98.57	99.53	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

Table A.8: All principal components and their variance for the Expt. 8 (physiographic variables and KG-rainfall spatial distribution derived from the three rain gauges under the R$_{assumed}$ scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	-0.07	**0.99**	0.05	0.02	-0.04	-0.13	0.06
Rbuffer	-0.44	-0.01	-0.32	**0.75**	-0.02	-0.06	-0.37
Surface slope	-0.11	-0.07	**0.93**	0.32	-0.14	0.00	-0.02
Elevation	-0.47	0.09	0.05	-0.30	0.05	**0.79**	-0.21
TWI	-0.03	0.02	0.13	0.06	**0.99**	-0.07	0.01
FL	**-0.52**	-0.09	-0.08	0.06	-0.02	-0.04	**0.84**
KG-rainfall	**0.54**	0.09	-0.09	0.49	0.04	**0.59**	0.33
Variance [%]	60.62	16.90	11.26	5.00	4.97	0.82	0.43
Cumulative variance [%]	60.62	77.52	88.78	93.78	98.75	99.57	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

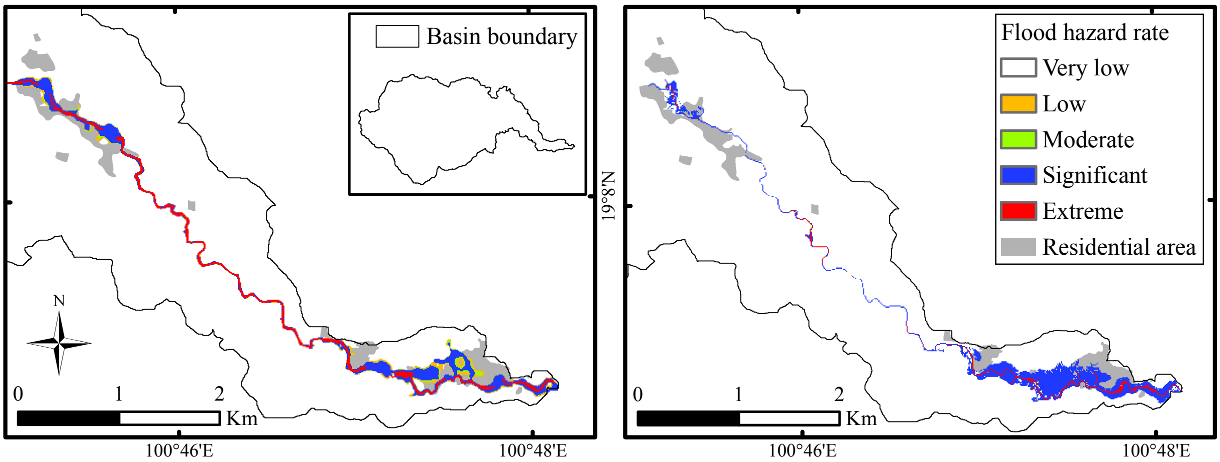
Table A.9: All principal components and their variance for the Expt. 9 (physiographic variables and TP-rainfall spatial distribution derived from the three rain gauges under the R$_{assumed}$ scenario)

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Dd	0.07	**0.98**	0.03	-0.15	0.00	0.02	0.11
Rbuffer	0.48	-0.02	-0.26	-0.08	-0.01	**0.77**	-0.31
Surface slope	0.11	-0.05	**0.95**	-0.09	-0.13	0.23	-0.02
Elevation	0.46	0.12	0.08	**0.54**	-0.09	-0.41	**-0.55**
TWI	0.03	0.02	0.14	0.20	**0.96**	0.07	0.04
FL	**0.52**	-0.07	-0.03	0.35	-0.12	0.00	**0.77**
TP-rainfall	**-0.52**	0.14	0.02	**0.71**	-0.17	0.41	0.02
Variance [%]	57.92	16.98	11.12	5.83	4.90	2.75	0.52
Cumulative variance [%]	57.92	74.90	86.01	91.84	96.74	99.48	100.00

Note: Absolute values of component coefficients greater than 0.50 are highlighted in bold.

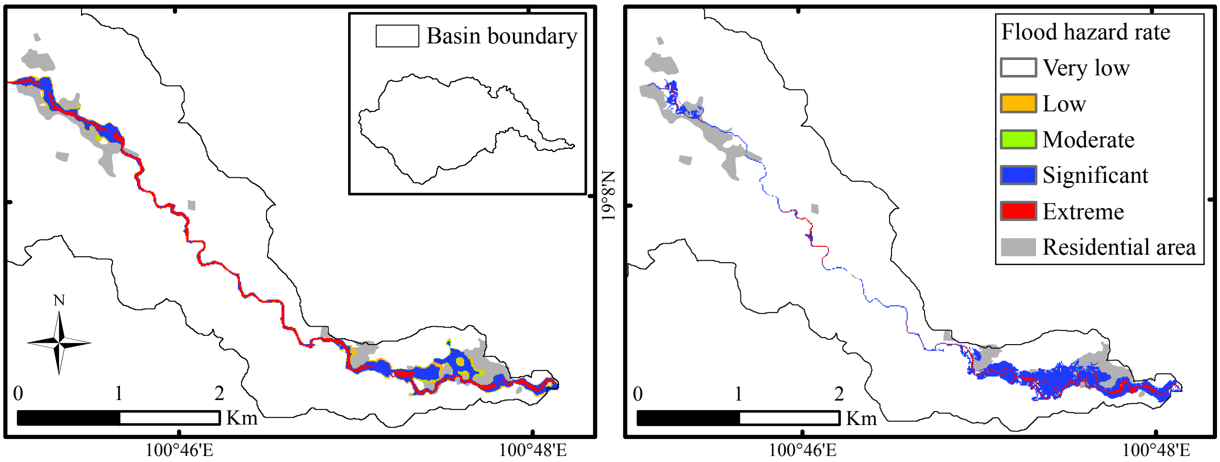
A.3.3 Flash flood hazard maps

A.3.3.1 Flash flood hazard maps produced by the integrated modeling and spatial index-based approaches



(a) (b)

Figure A.2: Flash flood hazard maps of the Nam Rim watershed (N1) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.



(a) (b)

Figure A.3: Flash flood hazard maps of the Nam Rim watershed (N1) under the regional rainfall scenario of RTr50: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

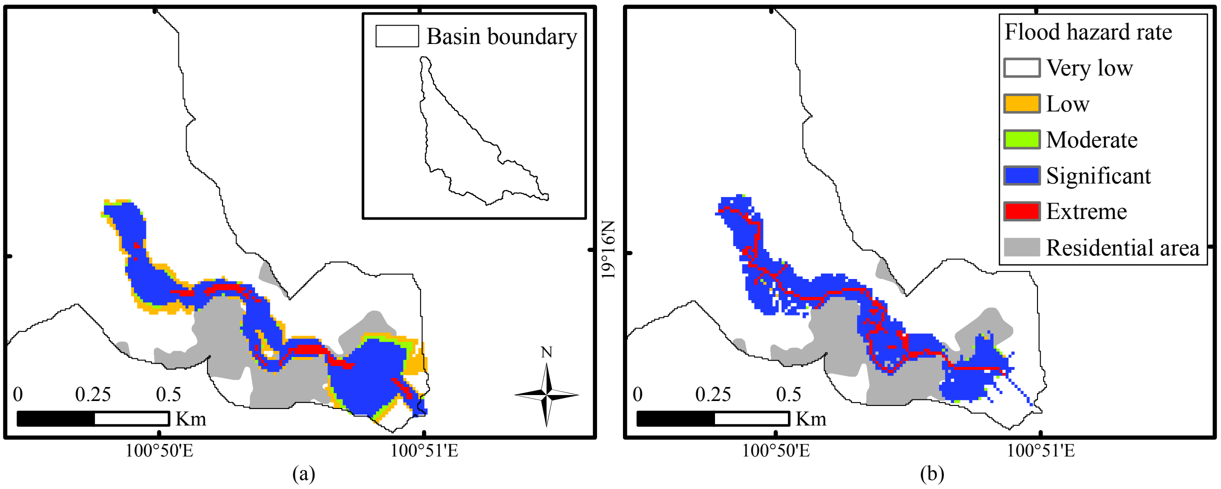


Figure A.4: Flash flood hazard maps of the Nam Hui watershed (N2) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

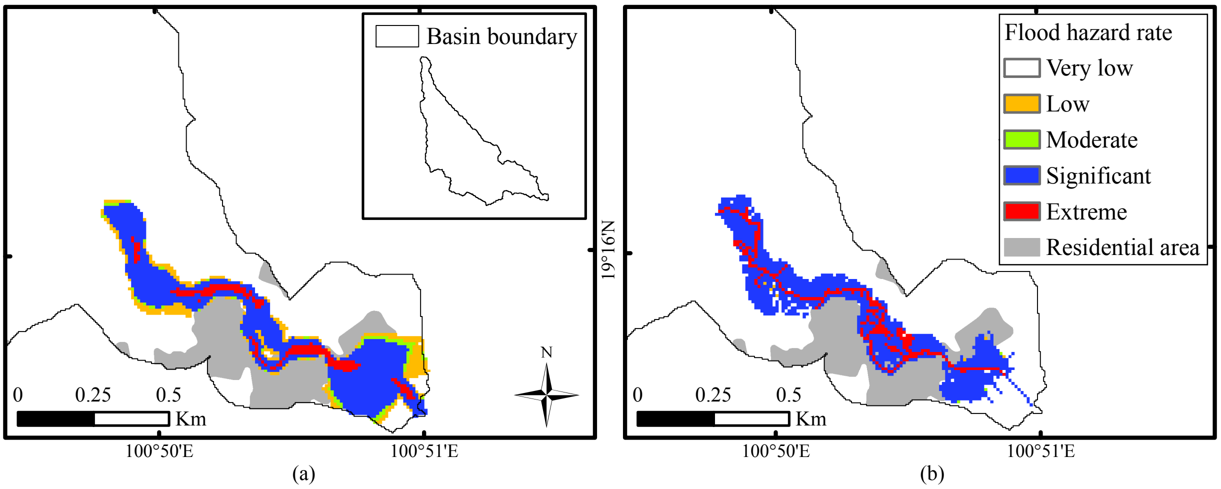


Figure A.5: Flash flood hazard maps of the Nam Hui watershed (N2) under the regional rainfall scenario of RTr50: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

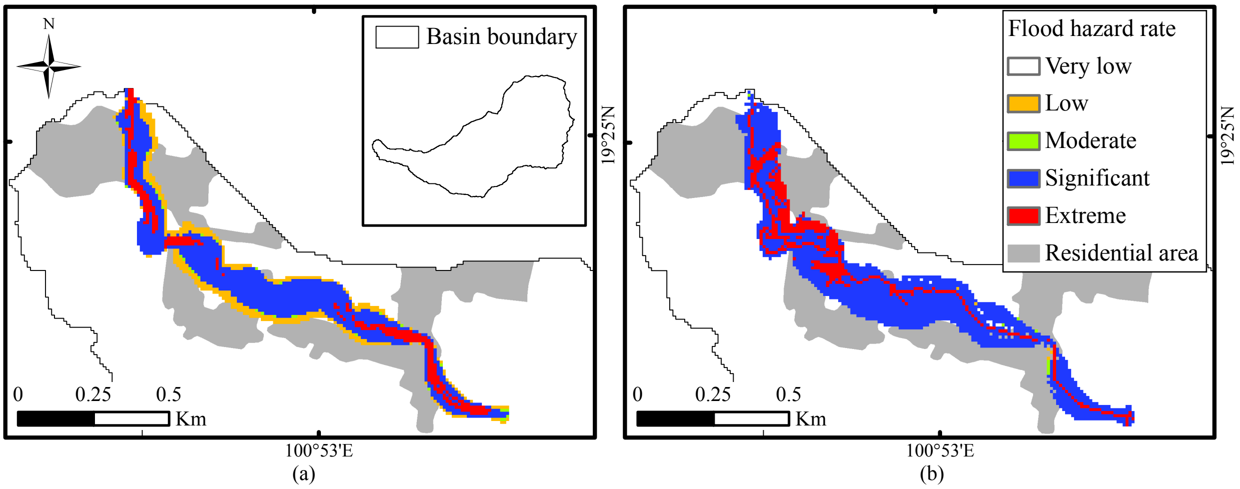


Figure A.6: Flash flood hazard maps of the Nam Lae watershed (N3) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

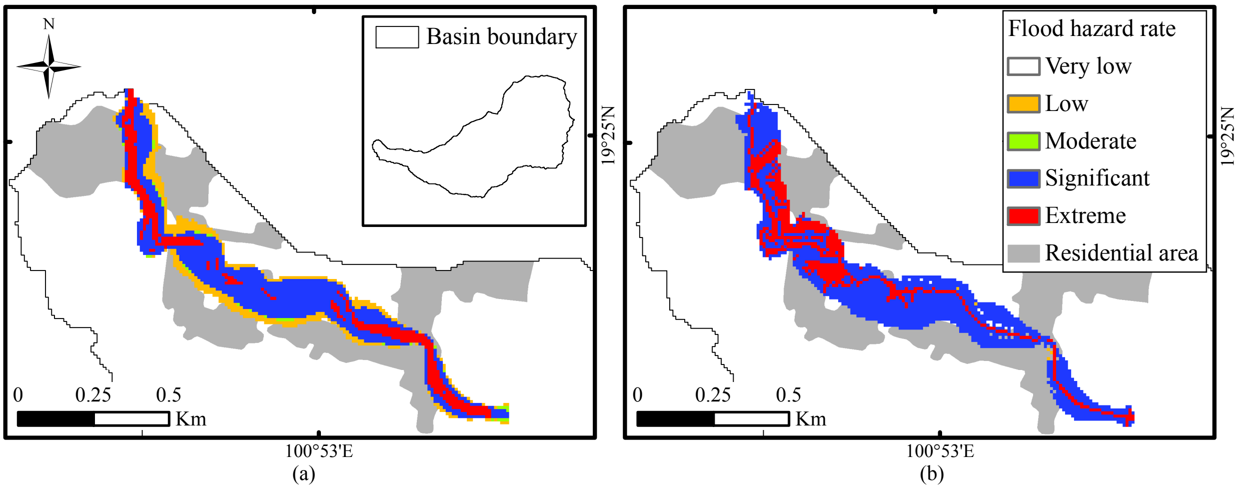
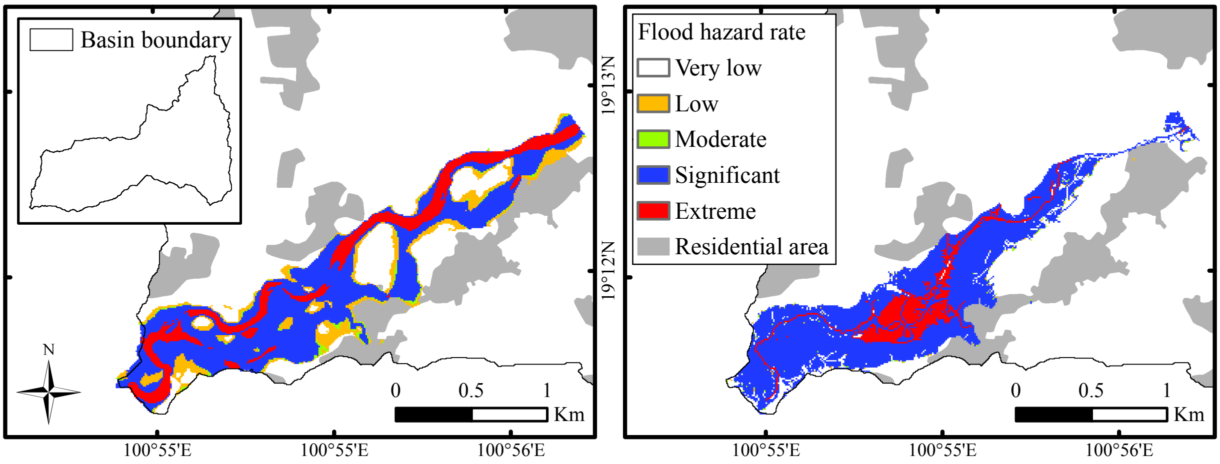
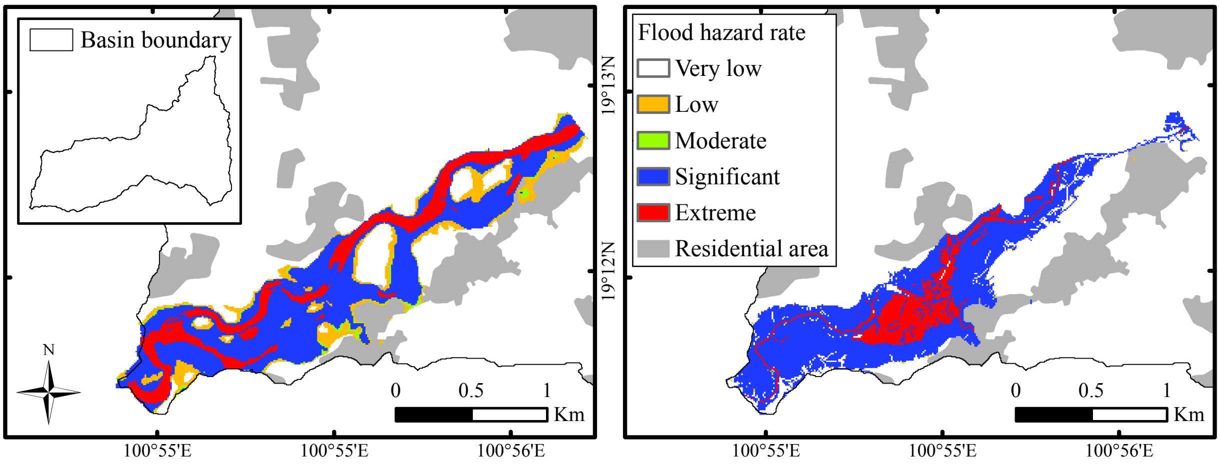


Figure A.7: Flash flood hazard maps of the Nam Lae watershed (N3) under the regional rainfall scenario of RTr50: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.



(a) (b)

Figure A.8: Flash flood hazard maps of the Nam Pua watershed (N4) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.



(a) (b)

Figure A.9: Flash flood hazard maps of the Nam Pua watershed (N4) under the regional rainfall scenario of RTr50: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

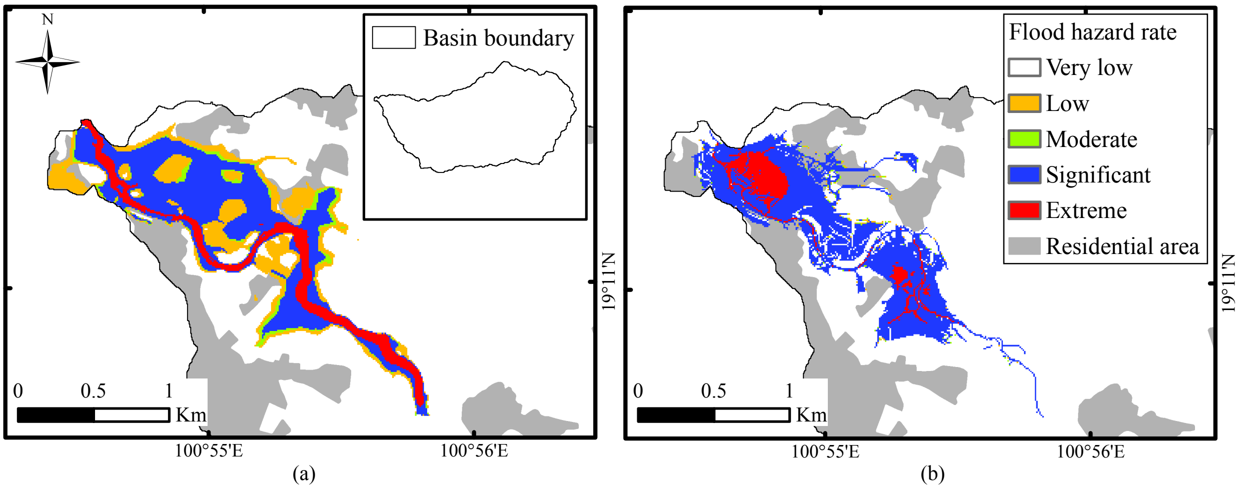


Figure A.10: Flash flood hazard maps of the Nam Khwang watershed (N5) under the regional rainfall scenario of RTr10: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

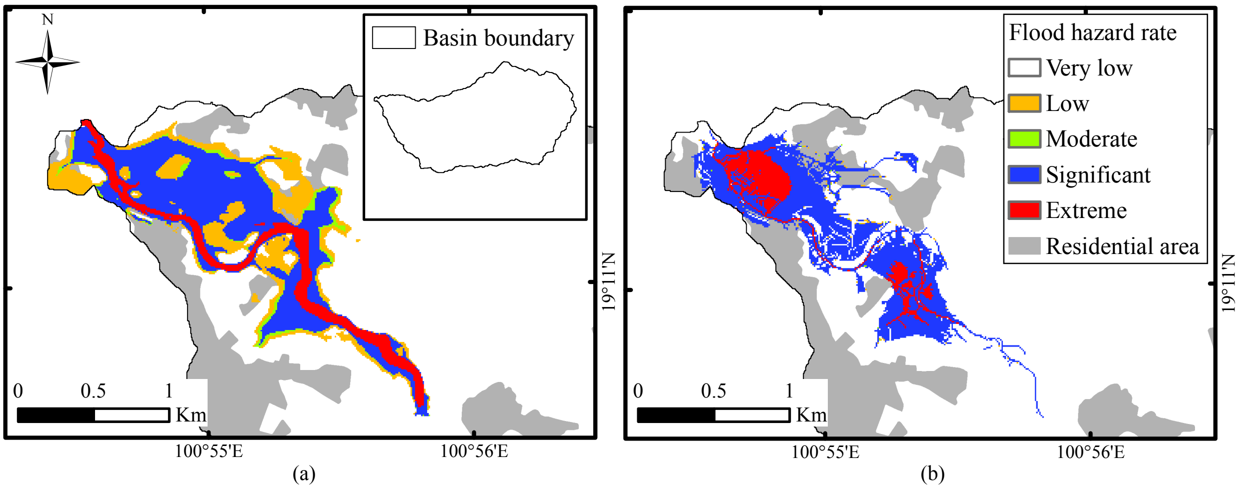
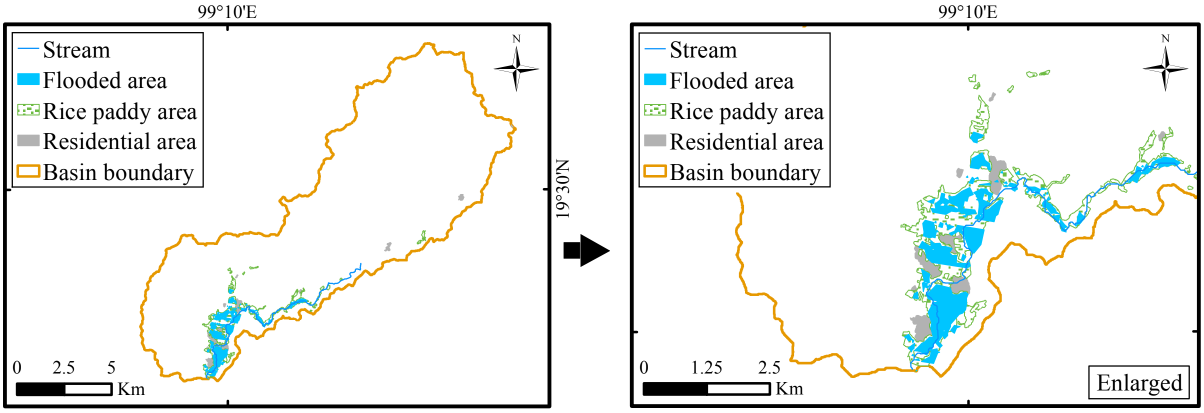


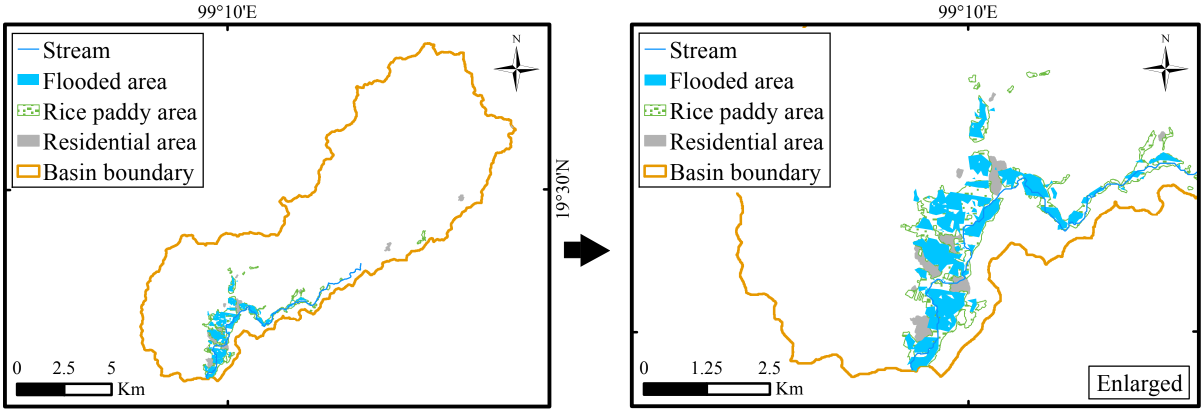
Figure A.11: Flash flood hazard maps of the Nam Khwang watershed (N5) under the regional rainfall scenario of RTr50: (a) produced by the integrated modeling approach and (b) produced by the spatial index-based approach.

A.3.3.2 Flash flood hazard maps in the upper Ping River basin



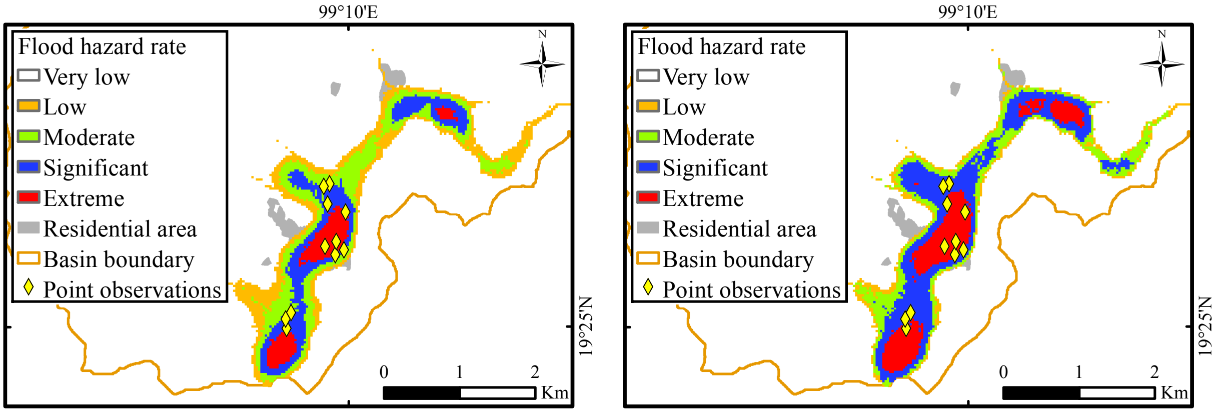
(a) Observed on 5 August 2003

(b) Observed on 5 August 2003 (with a closer view)



(c) Observed on 26 July 2011

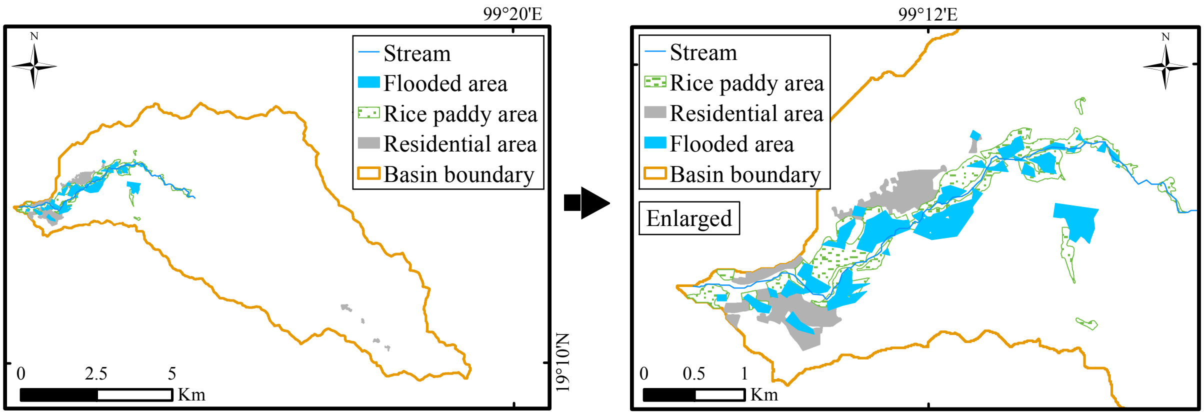
(d) Observed on 26 July 2011 (with a closer view)



(e) Natural breaks classification

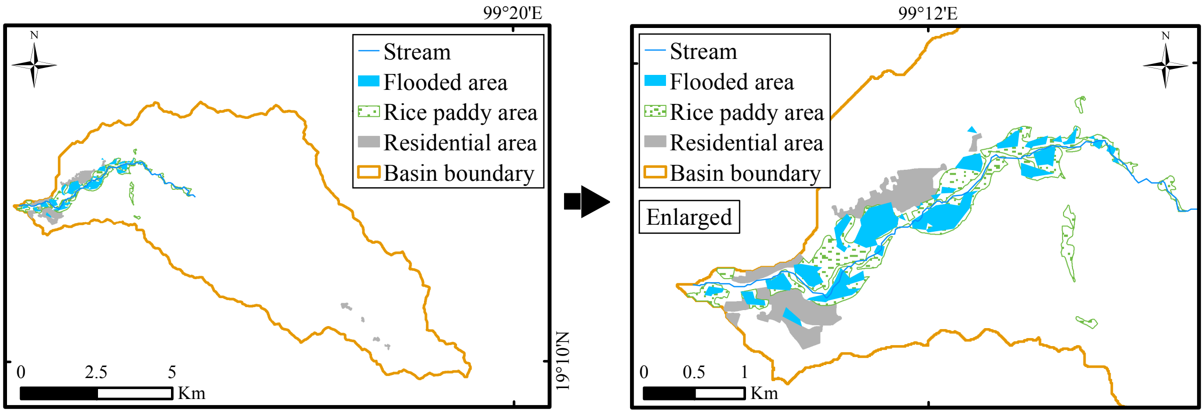
(f) Geometrical interval classification

Figure A.12: Inundated areas extracted from Landsat-7 in the Nam Mae Ngat watershed (P2): (a and b) Landsat-7 observed on 5 August 2003, (c and d) Landsat-7 observed on 26 July 2011, and flash flood hazard maps based on the spatial index-based approach using (e) natural breaks classification and (f) geometrical interval classification.



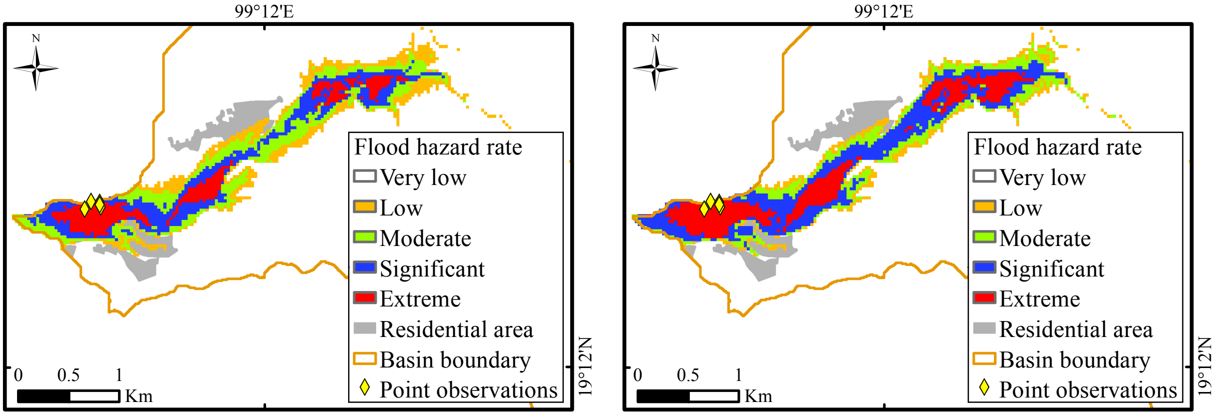
(a) Flooded areas extracted from Landsat-7 (on 5 August 2003)

(b) Flooded areas extracted from Landsat-7 (on 5 August 2003) with a closer view



(c) Flooded areas extracted from Landsat-7 (on 26 July 2011)

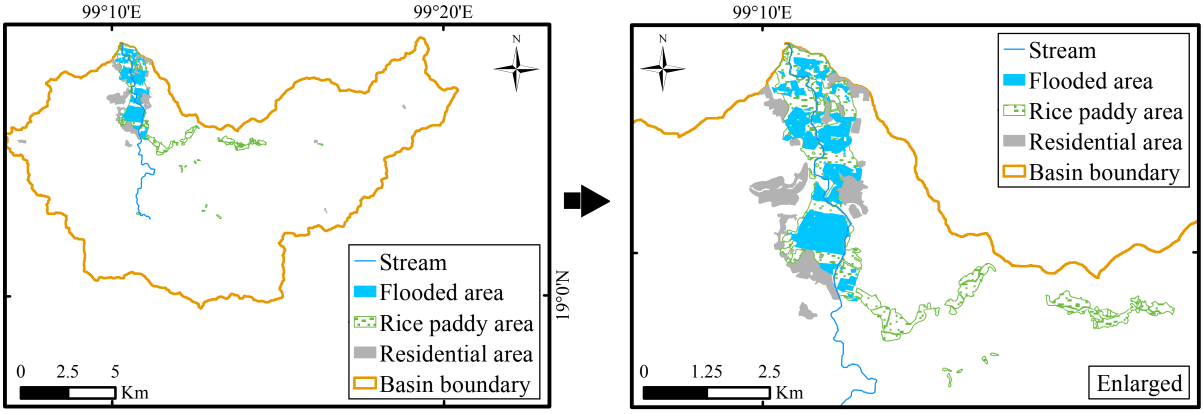
(d) Flooded areas extracted from Landsat-7 (on 26 July 2011) with a closer view



(e) Natural breaks classification

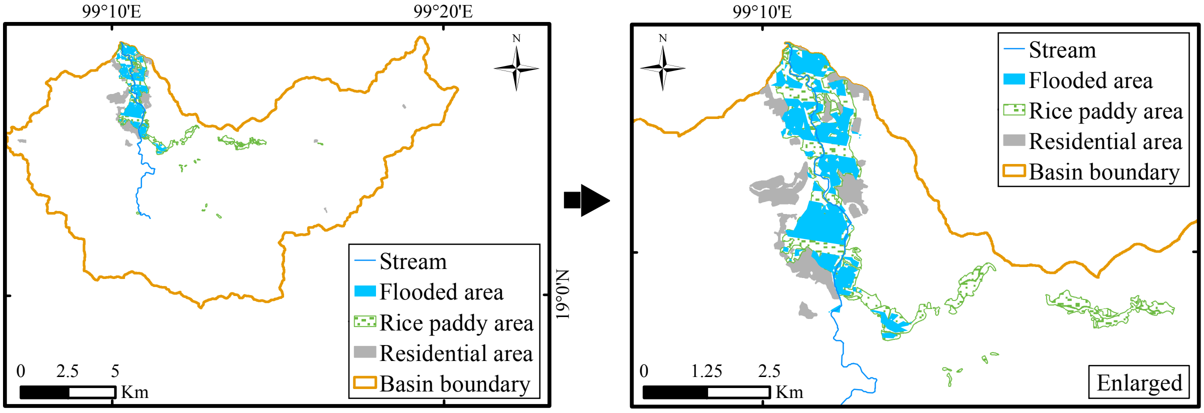
(f) Geometrical interval classification

Figure A.13: Inundated areas extracted from Landsat-7 in the Nam Mae Ngat branch watershed (P3): (a and b) Landsat-7 observed on 5 August 2003, (c and d) Landsat-7 observed on 26 July 2011, and flash flood hazard maps based on the spatial index-based approach using (e) natural breaks classification and (f) geometrical interval classification.



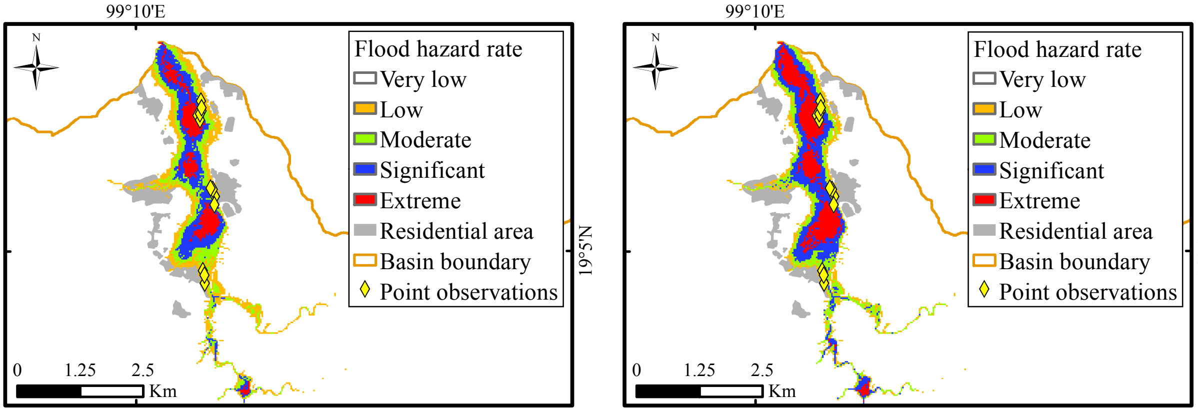
(a) Flooded areas extracted from Landsat-7 (on 5 August 2003)

(b) Flooded areas extracted from Landsat-7 (on 5 August 2003) with a closer view



(c) Flooded areas extracted from Landsat-7 (on 26 July 2011)

(d) Flooded areas extracted from Landsat-7 (on 26 July 2011) with a closer view



(e) Natural breaks classification

(f) Geometrical interval classification

Figure A.14: Inundated areas extracted from Landsat-7 in the Nam Mae Khot watershed (P4): (a and b) Landsat-7 observed on 5 August 2003, (c and d) Landsat-7 observed on 26 July 2011, and flash flood hazard maps based on the spatial index-based approach using (e) natural breaks classification and (f) geometrical interval classification.

[This page intentionally left blank]

B List of Abbreviations and Symbols

Abbreviations

AHP	analytic hierarchy process
AM24H	annual maximum 24 h
CMORPH	CPC MORPHing technique
CPC	Climate Prediction Center
cms	cubic meter per second
D	dimensional
d	day
DDF	depth-duration-frequency
DEM	digital elevation model
DMR	Department of Mineral Resources
DT	rule-based decision tree
DWR	Department of Water Resources
FFG	flash flood guidance
FFPI	flash flood potential indices
FHR	flood hazard rate
FN	false negatives
FP	false positives
FR	frequency ratio
GEV	generalized extreme value
GPA	generalized Pareto
HRU	hydrologic response unit
h	hour
IDW	inverse distance weighting
JAXA	Japan Aerospace Exploration
K	Kelvin
KG	kriging
kg	kilogram
km	kilometer
LDD	Land Development Department
LN	lognormal
mm	meter
min	minute
mm	millimeter
N	newton
NASA	National Aeronautics and Space Administration
PCA	principal component analysis
PC	principal component
PE3	Person type III
RFA	regional frequency analysis

RID	Royal Irrigation Department
SCS	Soil Conservation Service
TIN	triangular irregular network
TMD	Thai Meteorological Department
TP	true positives
TSP	Thiessen polygon
TRMM	Tropical Rainfall Measuring Mission
USDA	United States Department of Agriculture
USGS	United States Geological Survey
WoE	weights-of-evidence

Letter symbols

A	drainage area
A_s	slope of the saturated vapor pressure at mean air temperature
A_x	cross-sectional area of the flow
a	a parameter
a_{Fi}	acceleration due to Coriolis force
B_{bf}	channel top width at bankfull
b	a parameter
C	coefficient
C_{com}	compactness coefficient
C_s	coefficient
C_x	covariance matrix of variables
C_z	Chezy coefficient
CN	curve number
CN_{smd}	relative soil moisture curve number
c_f	friction coefficient
c_p	specific heat of the air at constant pressure
D	water depth
D_{bf}	hydraulic depth at bankfull
D_i	discordancy measure
Dd	drainage density
DF	debris factor
d	distance
E	evaporation rate
E_a	aerodynamic evaporation
E_r	evaporation rate based on net radiation
EI	Nash-Sutcliffe efficiency index
ET_p	potential evapotranspiration
e_a	vapor pressure
e_s	saturation vapor pressure
F	cumulative infiltration
F_{form}	form factor
F_{shape}	basin shape factor
F_{str}	stream frequency
FL	flow length
f_c	minimum infiltration capacity
f_h	infiltration capacity

f_0	initial infiltration capacity
G	Grubbs value
G	soil heat flux
g	acceleration of gravity
H	heterogeneity measure
H_b	basin relief
H_n	net radiation
HI	hypsometric integral
HQn	an n-year return period flood
h	water depth
I_a	initial abstraction
I_e	effective rainfall intensity
I_{slp}	basin slope index
K	storage constant
k	storage constant
k_{St}	Manning-Strickler coefficient
L_b	basin length
L_{lfp}	length of longest flow path
L_m	main channel length
L_o	length of overland flow
L_v	latent heat of vaporization
L-Ck	L-coefficient of kurtosis
L-Cs	L-coefficient of skewness
L-Cv	L-coefficient of variation
MAE	mean absolute error
m	number of items
N_{geo}	geometry number
N_p	number of items
N_{rug}	ruggedness number
N_{strord}	stream order number
n	number of items
P	precipitation
P_b	basin perimeter
P_e	depth of effiective precipitation
$PBIAS$	percent bias
Q	rate of discharge
Q_{bf}	bankfull flow
Q_e	equilibrium flow rate
Q_{in}	inflow rate
Q_{max}	maximum inflow rate
Q_{out}	outflow rate
Q_p	flooding flow
q_p	areal runoff
q_{pR}	unit hydrograph peak for a specific duration
q_{SM}	saturation of soil moisture
R	rainfall depth
R_{cir}	circulation ratio
R_{elong}	elongation ratio
R_f	infiltration ratio

R_i	rainfall intensity
R_{hy}	hydraulic radius
R_{tex}	texture ratio
R_{th}	runoff threshold
Rb	bifurcation ratio
Rbuffer	distance from stream
Rl	length ratio
RPE	relative peak error
$RMSE$	root mean square error
RTrn	an n-year return period rainfall event
r	rainfall depth
r^2	coefficient of determination
r_a	aerodynamic resistance
r_s	bulk surface resistance
S	volume of channel storage
S_f	friction slope
S_e	potential surface retention
S_m	mean basin slope
S_r	basin relief ratio
SM	relative soil moisture
T_{ns}	total number of streams
T_{sl}	total stream length
TWI	topographic wetness index
t	time
t_c	time of concentration or concentration time
t_r	effective rainfall duration
U	a matrix
u	velocity
V	velocity
v	velocity
X_m	normalized variable
X_{obs}	observed variable
X_{sim}	simulated variable
x	horizontal coordinate
x_m	value of variable m
y	horizontal coordinate
WL	water level
w	weight coefficient
Z	radar reflectivity
Z^{dist}	Z-statistic (a goodness-of-fit measure)
z	elevation
α (alpha)	Boussinesq coefficient
β (beta)	Horton’s decay parameter
γ (gamma)	psychrometric constant
θ (theta)	latitude
Λ (lambda)	a diagonal matrix
λ (lambda)	eivenvalue
λ_D	Darcy-Weisbach coefficient
ρ (rho)	air density

ρ_a	density of moist air
ρ_w	density of water
τ (tau)	unit shear
τ_{so}	unit shear at channel bed
τ_{wind}	unit shear due to wind
ω (omega)	angular velocity

Mathematical symbols

Σ	summation
$>$	greater than
$<$	less than
$\mid\ \mid$	absolute
Δ	an increment of
$\overline{X}$	the bar over a symbol denoting an average
∂	partial derivation
X^T	the power T of a symbol denoting a transpose operator of a matrix (symbol)